Validation of LAI, Chlorophyll and FVC biophysical estimates from Sentinel-2 Level 2 Prototype Processor over a heterogeneous savanna and grassland environment in South Africa

Philemon Tsele [1,*], Abel Ramoelo [2], Mcebisi Qabaqaba [2], Madodomzi Mafanya [1, 3], George Chirima [1, 4]

[1] *Department of Geography, Geoinformatics and Meteorology, University of Pretoria, Pretoria 0028, South Africa;*
[2] *Centre for Environmental Studies, Department of Geography, Geoinformatics and Meteorology, University of Pretoria, Pretoria 0028, South Africa;*
[3] *Department of Geography, University of South Africa, Pretoria, 0027, South Africa;*
[4] *Geoinformation Science Division, Agricultural Research Council Institute for Soil, Climate and Water, Pretoria 0001, South Africa; chirimaj@arc.agric.za*

*****Correspondence: philemon.tsele@up.ac.za; Tel.: +27-12-420-4939

The Sentinel-2 Level 2 Prototype Processor (SL2P) allows the generation of biophysical estimates at high spatiotemporal resolution from Sentinel-2 imagery and could be a solution for generating products in natural environments. This study validated the SL2P estimates of leaf area index (LAI), fractional vegetation cover (FVC) and canopy chlorophyll content (CCC) over the savanna and grassland environments using field measurements. The performance of the SL2P estimates in Marakele and Golden Gate Highlands National Parks were comparatively poor and linearly biased coupled with moderate-to-high errors. The SL2P estimates in the two study sites had low accuracy with relative root mean squared error's in the range 61.63% to 85.26% and possible systematic underestimations with pBias's ranging from 32.17% to 63.16%. These findings gave insight about the performance of the SL2P estimates over the considered heterogenous environments, and suggest the need for extensive validation and re-calibration of the system using long-term field measurements.

Keywords: leaf area Index (LAI); fractional vegetation cover (FVC); canopy chlorophyll content (CCC); error; Sentinel-2

Introduction

Estimation of vegetation biophysical variables is important for understanding vegetation condition and structure, growth status, nutritional stress, disease or insect infestation and gross primary productivity. In light of biodiversity loss and ecosystem restoration, these variables can be used to assess and monitor vegetation state and biodiversity at large. Climate change, invasive species, over-utilization of natural resources and land use change are cited as key causes of biodiversity loss (IPBES, 2019, GBO05, 2020, Living Planet Report, 2020, Climate change, 2021, GCOS, 2011). It is critical to continuously assess and monitor biodiversity to effectively and proactively plan its sustainable use and management.

The leaf area index (LAI) defined as the one-sided leaf area per unit of horizontal surface area (Jonckheere et al., 2004) is a key indicator of vegetation structure and growth, and also forms an essential input in climate models to determine ecosystem productivity, the surface energy balance and evapotranspiration (GCOS, 2011). Another biophysical variable called the leaf chlorophyll concentration (LCC), can be obtained through averaged SPAD (unit-less) leaf chlorophyll measurements. LCC carries valuable information about vegetation physiology and could be regarded as a key indicator of the actual plant health status. Accurate measurements of LCC can be helpful for precision management of natural resources and agricultural fields (Bei et al., 2019). Furthermore, the canopy chlorophyll concentration (CCC), which refers to the overall amount of chlorophyll *a* and *b* pigments in a compact group of plants per unit ground area (Gitelson et al., 2005) is derived from the product of the LCC, $\mu g.cm^{-2}$ and the corresponding LAI, $m^2.m^{-2}$ in a subplot (Darvishzadeh et al., 2008). CCC is an important indicator of

vegetation health condition, plant species diversity and forage quality assessment (Ali et al., 2020). Another key vegetation biophysical variable is called the fractional vegetation cover (FVC) and corresponds to the fraction of combined photosynthetic and non-photosynthetic vegetation separated from the exposed soil background within the total study area in the nadir direction. FVC is a key indicator of the spatial distribution of the vegetated area and the density of vegetation growth, and as a result plays a vital role in energy balance processes, development of terrestrial ecosystems and studying the interactions between climate change and vegetation (Wang et al., 2017, Liang and Wang, 2019). Therefore, the collection of vegetation biophysical properties such as the LCC, CCC, LAI and FVC provides critical information that could facilitate effective monitoring and management of natural vegetation at different spatial scales (Ali et al., 2020, Gitelson et al., 2006).

Natural heterogeneous canopies like the grasslands of South Africa, are characterized by native grasses of different mixture of species in varying proportions, often distributed across fluctuating terrain slopes (Masemola et al., 2016, Ramoelo et al., 2015). Such an environment, analogous to rangelands, is favourable for livestock production and grazing by animals (Svinurai et al., 2021). Therefore, it is critical to (i) assess areas where there is a change in response to climate and/or anthropogenic effects, (ii) quantify the amount of aboveground biomass and vegetation cover, and (iii) monitor the functional status and diversity of the rangeland vegetation communities in-order to enhance ecosystem productivity and stability, guided by effective resource management strategies and policies.

These aforementioned processes are measurable through indicators such as LAI, LCC, CCC and FVC which can be estimated in the field or through modelling procedures applied to satellite remote sensing observations. Estimates of biophysical variables

acquired through field campaigns have proven to be reliable and are often used as authentic inputs in model calibration and validation studies (Lv et al., 2021, Brown et al., 2021, Hu et al., 2020, Kganyago et al., 2020). However, field-based estimates are expensive, time consuming and laborious to acquire especially for large geographic areas over a long-term temporal period. On the other hand, satellite remote sensing provides an alternative approach in estimating these biophysical variables over broad spatial extents on a regular basis, spanning a long period of time (Myneni et al., 2002). These four variables are also listed and ranked as some of the top 30 biodiversity metrics measured from space, using satellite remote sensing (Skidmore et al., 2021).

A number of available global biophysical products are generated from coarse to moderate spatial resolution satellite sensors such as, Advanced Very High Resolution Radiometer (AVHRR) (García-Haro et al., 2018), Moderate Resolution Imaging Spectroradiometer (MODIS) (Jia et al., 2018, Disney et al., 2016), PROBA-Vegetation (Baret et al., 2013), ENVISAT Medium Resolution Imaging Spectrometer (MERIS) (Dash and Curran, 2004, Bacour et al., 2006) on a regular basis over different time periods. In addition, given the lapsed ENVISAT mission for nearly 10 years, the Sentinel-3 Ocean and Land Colour Instrument (OLCI) provides continuity in the generation of global biophysical products (Pastor-Guzman et al., 2020). These biophysical products provide modelled estimates of for example LAI, FVC and CCC; and are widely used for monitoring and modelling of agricultural activity (Kganyago et al., 2020) and ecosystem dynamics (Hill et al., 2006, Cho et al., 2014). Notwithstanding the utility of these products particularly at regional to global scales, their spatial resolutions, which range from approximately 300 m to 1000 m could make it difficult for the products to provide reliable estimations of vegetation biophysical properties in heterogenous landscapes on a local scale (Lv et al., 2021).

In recent times, there has been a successful attempt to generate vegetation biophysical products that provide modelled estimates of LAI, FVC, CCC and Fraction of Absorbed Photosynthetically Active Radiation (FAPAR) at an increased spatial resolution of up to 10 m. In particular, Weiss and Baret (2020) developed the Sentinel-2 Level 2 Prototype Processor (SL2P) tool using a combination of artificial neural networks (ANNs) and physically-based modelling methods. The SL2P tool is integrated in the Sentinel Application Platform (SNAP) software version 8.0 (https://step.esa.int/main/download/snap-download/) and is freely accessible to users to generate biophysical products from atmospherically-corrected Sentinel-2 multi-spectral instrument (MSI) data. Sentinel-2 MSI data has 13 spectral bands, characterised by fine spatial resolutions i.e. 10-60 m and a high temporal resolution of up-to 5 days (Frampton et al., 2013). Furthermore, Sentinel-2 data is freely available, cover large spatial extents (i.e. 290 km x 290 km per scene) regularly throughout all seasons and as a result, has the potential for improved estimation of vegetation biophysical variables on a local scale (Hu et al., 2020, Hauser et al., 2021, Darvishzadeh et al., 2019, Ramoelo and Cho, 2018).

There have been recent numerous attempts to validate the SL2P derived biophysical variables over different vegetation types in various parts of the globe. For example Xie et al. (2019) validated the SNAP SL2P-derived LAI and CCC biophysical variables over winter wheat in Beijing, China and found the SL2P being applicable for accurate winter wheat crop LAI and CCC estimation. Brown et al. (2019) assessed the accuracy of the SL2P-derived LAI and CCC biophysical variables over a deciduous broadleaf forest site in Southern England. In particular, the study found that the aforementioned SL2P biophysical variables revealed moderate inaccuracies, and suggested using alternative model inversion methods such as the invertible forest reflectance model (INFORM) that are optimised for forest environments. Furthermore,

Djamai et al. (2019) validated three SL2P biophysical variables i.e. LAI, FVC and canopy water content (CWC) derived from Sentinel-2 MSI and Landsat-8 Operational Land Imager (OLI) imagery over different crop types such as soybeans, wheat, canola, oats black beans and alfalfa located in Manitoba, Canada. The SL2P showed satisfactory performances in estimating LAI and FVC (except for CWC which is a matter for further investigation) across all crop types on both MSI and OLI imagery. A virtually similar study was undertaken by Filipponi (2021) who compared SL2P-derived LAI estimates from Sentinel-2 MSI and Landsat-8 OLI images over croplands, grasslands, broadleaved and needleleaf forests in Italy. The study reported a general underestimation of LAI over the aforementioned test site classes, however higher overestimations were noted for the grasslands.

However, the opposite result was reported in Brown et al. (2021) whereby, an extensive validation of the SL2P LAI estimates against numerous ground samples of LAI over homogenous canopies comprising for example, grasslands and crops revealed overall satisfactory performance compared to in forest canopies, where the SL2P LAI retrievals experienced underestimation. Hu et al. (2020) evaluated the SL2P-derived LAI, FAPAR and FVC estimates using ground measurements of different vegetation types located in five countries and found that, the *Bias* and root mean squared error (*RMSE*) values of all SL2P biophysical estimates for forests and grasses were better than for crops. These findings were limited to a relatively small number of ground measurements especially over grasses (8 samples) and forests (19 to 20 samples) compared to crops with samples that ranged from 70 to 100. A more recent study in South Africa by Kganyago et al. (2020) validated the SL2P-derived LAI estimates over sunflower and maize crop types using ground LAI measurements. Notwithstanding the reported moderate agreements (coefficient of determination, R^2 of about 0.70), the computed errors (*Bias*

and *RMSE*) of the SL2P LAI estimates were found to be relatively high in an African semi-arid agricultural landscape.

While some of the SL2P biophysical variables have demonstrated satisfactory performance over agricultural environments (Hu et al., 2020, Kganyago et al., 2020) the performance of different SL2P derived biophysical estimates particularly over natural heterogenous landscapes characterised by diversity of land cover, species diversity and varying terrain slopes remains largely untested (Brown et al., 2021). Moreover, there is a need to broaden the assessment of errors in different biophysical estimates, derived from the SL2P tool across different regions, vegetation types and temporal periods. This exercise may not only be critical for identifying the sources of error in deriving these biophysical estimates, but also in providing better insights on their performances.

As a result, the aim of this study was to validate three SL2P-derived biophysical estimates over a heterogenous savanna and grassland environment in South Africa using field measurements of grass biophysical variables. Our key question is, could the SL2P-derived biophysical parameters capture the variability of the vegetation state in the study area and significantly relate to field ground measurements? The specific objectives of this study were to: (i) quantify the error of SL2P-derived LAI, CCC, and FVC using ground measurements during the wet season in the grassland areas of two South African national parks i.e. Golden Gate Highland National Park (GGHNP) and Marakele National Park (MNP); (ii) Compare the reliability and consistency of the three SL2P-derived biophysical estimates in the grassland areas of GGHNP and MNP.

Materials and methods

Study Area

Two heterogenous study sites were selected in two South African National Parks namely, the GGHNP located in the Free State province near the Lesotho border, and MNP located in the Waterberg mountains of the Limpopo province, as shown in Figure 1. The study sites were selected based on key location attributes, which encompassed the savanna and grassland biomes and different vegetation communities, according to the national vegetation map (Mucina and Rutherford, 2006). Furthermore, the GGHNP and MNP sites are mountainous and characterized by surface height variation i.e. elevations that range between approximately 1639 m to 2815 m and 976 m to 2091 m respectively, estimated from the 30 m resolution Shuttle Radar Topography Mission (SRTM) data acquired from the United States Geological Survey (USGS) Earth Explorer (https://earthexplorer.usgs.gov/). Both sites fall within the summer rainfall region of South Africa. In particular, the GGHNP receives average rainfall of approximately 700 mm per year (Mucina and Rutherford, 2006) whereas, the MNP site can receive average rainfall of up to around 630 mm annually (Van Staden and Bredenkamp, 2005).

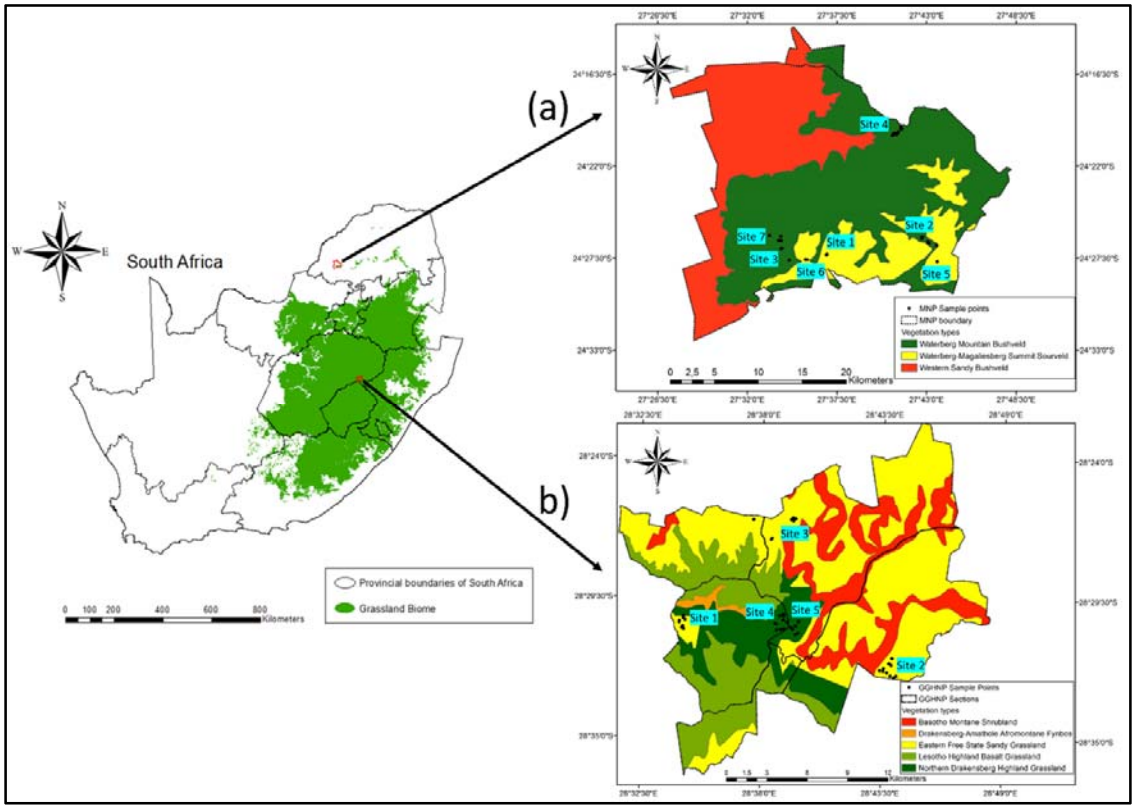

Figure 1. The location of the two selected study sites in South Africa (a) and (b) where LAI, CCC and FVC sample field measurements were taken. Site (a) represents the Marakele National Park (MNP) whereas site (b) represents the Golden Gate Highlands National Park (GGHNP) on 8 – 10 April 2021 and 18 – 21 March 2021 respectively.

The dominant underlying geology in the GGHNP includes mudstone, fine-to-medium sandstone and basalt, based on the national geology map developed by the Council for Geosciences (CG, 1997). In the same site, the soils include shallow to deep sandy soil that is extremely gravelly as well as a clay-rich subsoil (https://data.isric.org/; Van Engelen and Dijkshoorn, 2013). On the other hand, in MNP the geology is largely characteristic of sandstone and mudstone, followed by granule stone, siltstone and diabase (CG, 1997). The soils in MNP range from shallow-gravel soil to low activity clayed soil (https://data.isric.org/; Van Engelen and Dijkshoorn, 2013).

Field data collection

Field data collection in the two study sites (Figure 1) took place during the wet season of 2021, during peak productivity for the natural heterogenous grasses (Masemola et al., 2016). The total number of sampled locations were 80 and 68 in GGHNP and MNP, respectively. The sampling strategy involved a combination of stratified and purposive sampling methods (Lv et al., 2021). Random samples were initially taken across different grass vegetation communities (Figure 1) and varying slope terrains spanning the crests, valleys and low to mid-slopes. However, when in the field, there were certain inaccessible areas, which led to the use of purposive sampling where re-placement of the sampled locations was done, close to the randomized points. Each selected sample location represented a plot with a size of 20 m x 20 m and within that plot, two subplots of size 1 m x 1 m spaced apart were taken in-order to capture variability within each plot.

A number of recordings were taken in each subplot namely, the (i) subplot number and photo (ii) geographic coordinates using the Global Navigation Satellite System – Real Time Kinematic (GNSS-RTK) method (Schloderer et al., 2011), (iii) LAI using the ACCUPAR LP-80 ceptometer, (iii) LCC using the SPAD 502 Plus chlorophyll meter, (iv) Grass height (cm) using the disk pasture meter and (v) observed proportions of photosynthetic vegetation (PV), non-photosynthetic vegetation (nPV) and bare soil (BS) to determine FVC. Field data collection was done on 18 – 21 March 2021 and 8 – 10 April 2021 in GGHNP and MNP, respectively. On these dates, the weather conditions were favourable and fairly stable characterised by sunny, warm temperatures and non-windy conditions. Furthermore, the dates falls within the mid- to late wet season, and as a result, the grasses in the study sites mostly appeared green and healthy (Figure 2).

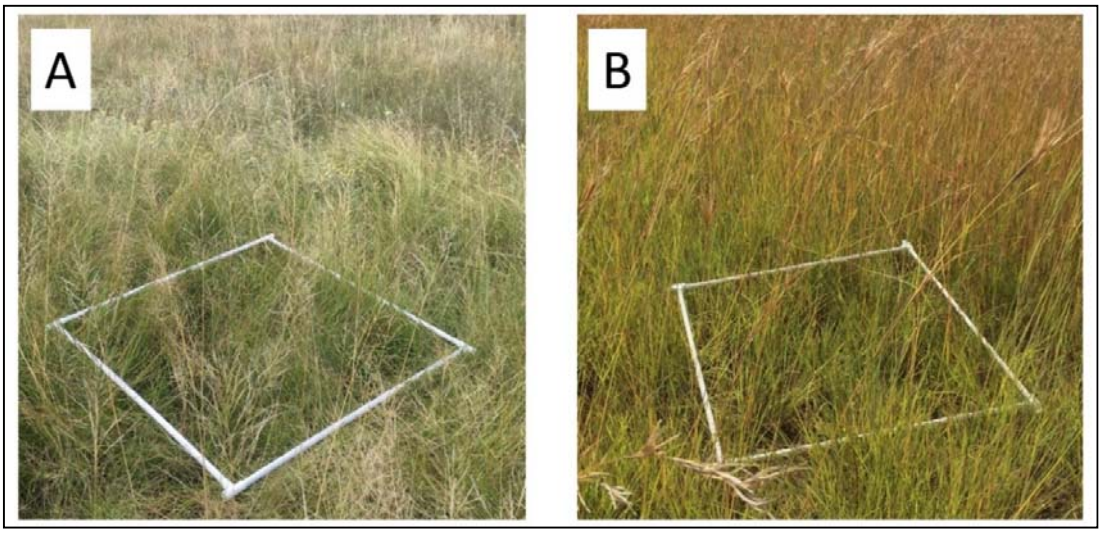

Figure 2. Example of the 1m x 1m subplots in GGHNP (A) and MNP (B). Within each subplot, the field estimates of grass biophysical parameters were taken.

Description of field data collection procedures

The two sampled 1 m x 1 m subplots within the Sentinel-2 20 m x 20 m pixel were considered sufficient to characterise the heterogeneity of the cover within the pixel. The chosen park sections in both GGHNP and MNP where the subplots were sampled, were homogenous in-terms of vegetation cover and slope (either mid-slope, valley or crest) (see Figure 3). It is this homogeneity which gave an indication that having two subplots within a 20 m x 20 m pixel is representative of the heterogeneity of the pixel cover in our study sites, because the vegetation cover appears less variable (Figure 3). This does not rule out the fact that in these surroundings there are several (more than 60) grass species present, and not all the sections may have a homogenous cover.

Figure 3. Example of the surrounding landscape in Golden Gate Highland National Park (GGHNP) where the subplots were sampled, characterised by a homogenous vegetation cover.

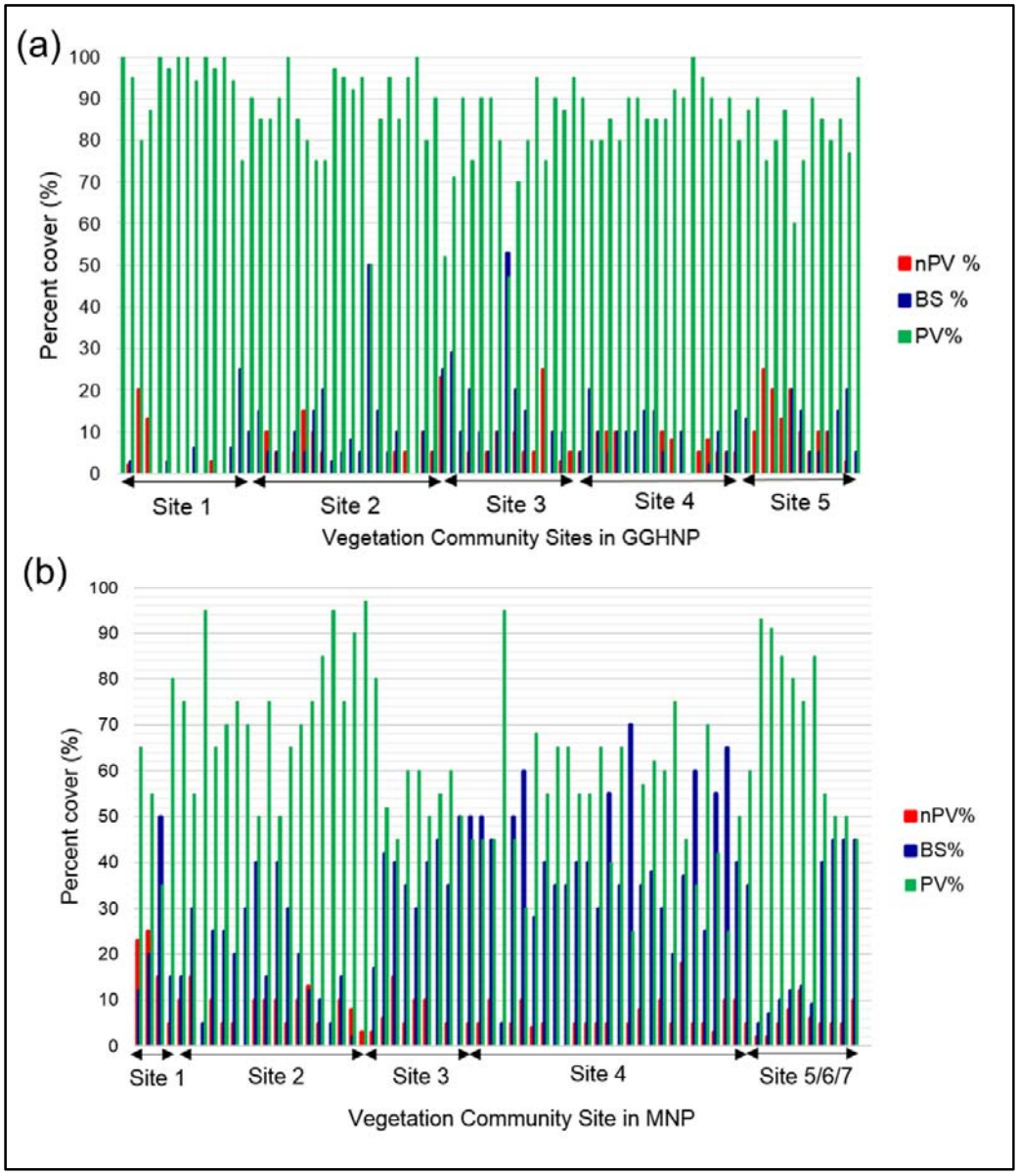

Figure 4. The percent cover (%) distribution of the three classes, namely non-photosynthetic vegetation (nPV), bare soil (BS) and photosynthetic vegetation (PV) across the subplots sampled in different sections or sites of the Golden Gate Highland National Park (GGHNP: (a)) and Marakele National Park (MNP: (b)).

To confirm the greenness of our subplots, visual inspection was conducted in the field by observing (in the nadir direction) the fractional covers of BS%, PV% and nPV% within each of the subplots. In particular, in each subplot, the fractional covers were estimated based on visual interpretation by three participating experts in the field, and thereafter an average was recorded per cover. For example, Figure 4 show the percentage

(%) distribution of the three covers i.e. BS, nPV and PV across the subplots located different sections of the parks. It is clear in GGHNP that PV (corresponding to green plant material) was highly dominant (greater than 70%) in most of the sampled subplots, whereas nPV (corresponding to brown plant material) was mostly below 10% (Figure 4 (a)). The MNP sites had on average about 10% of nPV and appeared similar to the overall nPV across the GGHNP sites (Figure 4 (b)). Therefore, this analysis was important to ensure that the presence of the brown material within the subplots is minimal and to some extent had a negligible influence on the measurements of LAI using the ACCUPAR LP-80 ceptometer (hereafter, referred to as ACCUPAR). Furthermore, the highly dominating green vegetation cover within our subplots, confirm that the period chosen for field data collection was peak productivity for the natural grasses in our study area.

Furthermore, in each subplot, LAI computations were performed using the ACCUPAR which measures photosynthetically active radiation (PAR; 400-700 nm) using 80 individual sensors distributed on its probe (https://www.metergroup.com/environment/products/accupar-lp-80-leaf-area-index/, Fang et al., 2014, Francone et al., 2014, Confalonieri et al., 2014). In particular, the ACCUPAR assumes the leaves in a canopy volume are randomly distributed (Chen et al., 2005), and therefore it is insensitive to leaf distribution variability when measuring LAI (Fang et al., 2014). In particular, the ACCUPAR measurements of LAI in this study are considered 'effective' because the ACCUPAR instrument partially accounts for foliage clumping using the logarithm averaging method (Lang and Yueqin, 1986), and therefore measures effective LAI, denoted LAI_e. Throughout the paper, LAI is used interchangeably with LAI_e. This instrument takes into cognisance the leaf angle distribution parameter χ and gap-fraction analysis when computing LAI by measuring incoming PAR above and below the vegetation canopy (Campbell, 1986). Although

different ranges of parameter χ values are provided in the manual to describe canopy angle distributions (mostly for crops), in this study the χ parameter of the ACCUPAR was kept to the default value of 1 meaning a spherical leaf angle distribution is assumed, due to the heterogenous nature of the grass species across the sampled plots in GGHNP and MNP. For each subplot, the incident PAR above the grass subplot was measured, followed by measurements of below-grass-canopy PAR in-order to calculate the average LAI. The assumption with the calculations of LAI using the ACCUPAR is equivocalness, in addition to other contributing factors such as foliage clumping, row spacing and leaf distribution variability (Johnson et al., 2010). In this study, LAI readings were performed under generally clear skies with intermittent cloud cover from the late morning hours at about 10:00 until early afternoon at around 14:00 in-order to minify variations of the sun zenith angle among the subplots.

Moreover, in each subplot we used the SPAD 502 Plus chlorophyll meter to take unit-less chlorophyll readings of five randomly selected green leaves, representing the dominant species and recorded the average chlorophyll reading. The average chlorophyll readings (i.e. SPAD measurements) of all subplots were converted into LCC per unit area, $\mu g.cm^{-2}$ by applying an empirical calibration method described in Markwell et al. (1995). The transfer formula from SPAD to LCC is defined in Equation (1) as:

$$LCC\ (\mu g.cm^{-2}) = 0.0893*(10^{SPAD\hat{}0.265}) \quad (1)$$

Subsequently, the LCC was multiplied by our field measured LAI to obtain the CCC in all subplots across the study sites. Lastly, the readings of grass height (cm) using the disk pasture meter were taken per subplot. For each subplot, this entailed taking one reading at the center of the subplot and four random readings around the perimeter of the subplot. All five readings were averaged into a single reading of the subplot grass height,

which was considered representative of the grass height variability within and around each subplot.

Schematic workflow

Figure 5 show a schematic workflow summarizing the various phases of the methodology that were implemented in this study. These phases are discussed in subsequent sections of the paper.

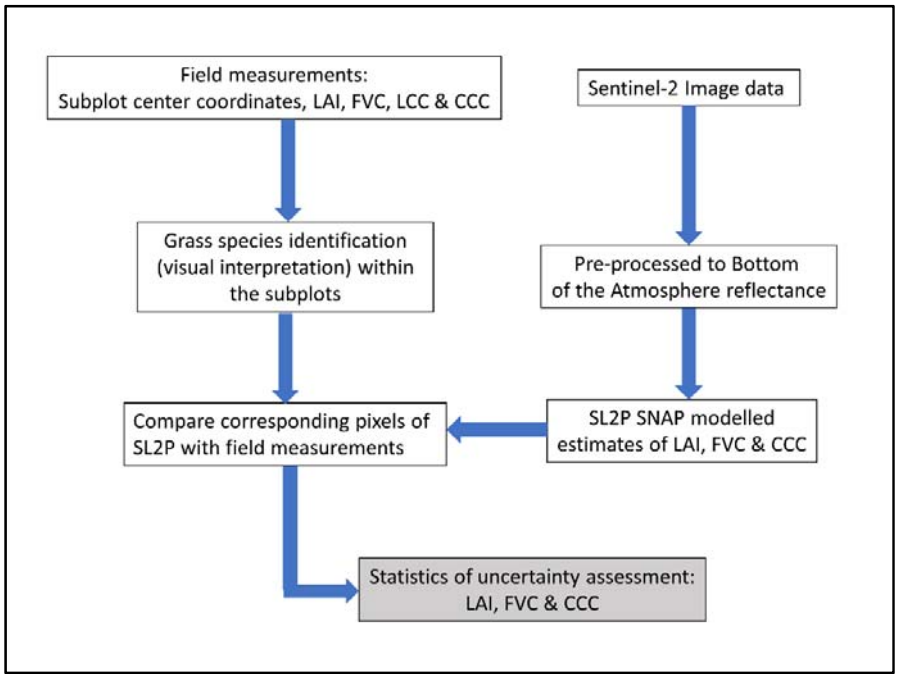

Figure 5. Summary of the methodology to validate the biophysical estimates of leaf area index, chlorophyll and fractional vegetation cover from Sentinel-2 Level 2 Prototype Processor over a heterogeneous savanna and grassland environment in South Africa.

Description of data-sets used

Remotely-sensed imagery

Sentinel-2 images were acquired free of charge from the European Space Agency data hub (https://scihub.copernicus.eu/dhus/#/home) on the 27th of March 2021 and the 9th of April 2021. The selection of the images was such that they (i) are free from any cloud obscuration (ii) covered the two study sites and (iii) had acquisition dates that were very close (i.e. <= 6 days) to the field data collection dates. The Sentinel-2 images were pre-

processed to surface reflectance or Bottom of the Atmosphere (BOA) reflectance i.e. Level-2A using the SNAP Sentinel-2 atmospheric correction tool, Sen2Cor, version 2.8 (Louis et al., 2016). Furthermore, the Sentinel-2 BOA images were resampled to the spatial resolution of 20 m using the resampling function within the SNAP software. This spatial resolution is such that the 20 m x 20 m pixels in sampled geographic areas correspond to single field plots of size of 20 m x 20 m that contains two subplots, each of size 1 m x 1 m described earlier. The SNAP SL2P tool was used in this study to generate the LAI, CCC and FVC grass biophysical variables using the resampled, atmospherically corrected Sentinel-2 reflectance bands. According to the design of the SL2P tool, the 20 m resolution biophysical products are generated using only eight reflectance bands of Sentinel-2 data (Weiss and Baret, 2020). Lastly, the geographical coordinates of the field sublots were used to extract corresponding Sentinel-2 20 m resolution pixels of LAI, CCC and FVC modelled estimates in both GGHNP and MNP. The extracted pixel values were validated using the field biophysical estimates.

SL2P-derived biophysical variables

The SNAP SL2P tool developed by Weiss and Baret (2020) uses an assortment of backpropagation ANNs that are trained with radiative transfer model (RTM) simulations. The simulations of the training database takes into cognisance the leaf optical properties, canopy structure, soil background reflectance and acquisition geometry parameters. Simulations were performed on a single joint probability function of canopy variables (designed to reflect global conditions) over natural vegetation and cropland areas. In particular, during simulations, the Leaf Optical Properties Spectra (PROSPECT) (Feret et al., 2008) and Scattering by Inclined Leaves (SAILH) (Verhoef, 1984) models were used to provide retrievals and predicted errors of CWC, LAI, CCC, FVC and FAPAR biophysical variables. A key assumption of the embedded algorithm within the SL2P tool

is that any target pixel is located on a landscape patch that is homogeneous enough in-order to avoid random fluctuations of radiation fluxes. However, in the case where the target pixel exhibits strong heterogeneity, the algorithm adds a flag to such a pixel, because it violates the assumption of the RTM used to build the neural network and possibly, the values of the variables used to train the database. This could result in a loss of accuracy in retrieving or estimating the biophysical variables (Weiss and Baret, (2020). One of the main advantages of the SL2P algorithm is that, it is generic without inputs of specific land cover types which could facilitate efforts to estimate vegetation biophysical variables at regional to global scales. In this study, the SL2P tool was used to retrieve the LAI, CCC and FVC biophysical variables using eight pre-processed Sentinel-2 bands over the GGHNP and MNP sites. There are three quality indicators (QA) provided by the SL2P which labels the quality of the retrievals (Weiss and Baret, 2020). In our study, the LAI, CCC and FVC biophysical variables were categorised as best retrievals by the SL2P algorithm (i.e. QA = 0).

Field measurements of biophysical variables

Table 1 show the summary statistics of the grass biophysical variables and terrain attributes over the two study sites. Generally, the field measurements across the subplots, resembled an approximately normal distribution, which was inferred from the proximity of the respective mean and median values per variable. The difference between these two basic statistical measures i.e. measures of central tendency, was considered in this study as a natural test for data distribution symmetry (Gastwirth, 1971).

Table 1. Summary statistics of selected terrain attributes and measured biophysical variables of grassland sample subplots. The statistical parameters, CV denotes the coefficient of variation, and StDev the standard deviation.

Site	Measured variables	No. of Subplots	Min.	Max.	Mean	Median	StDev	CV
GGHNP	LAI ($m^2.m^{-2}$)	80	0.61	6.24	2.26	2.02	1.24	0.55
	CCC ($\mu g.cm^{-2}$)	80	7.24	162.61	46.01	37.05	32.26	0.70
	FVC	80	0.47	1.00	0.86	0.87	0.11	0.13
	Grass height (cm)	80	5.00	34.00	11.92	11.00	5.48	0.46
	Elevation (m)	80	1832.20	2102.41	1966.04	1960.56	78.36	0.04
	Slope (°)	80	0.49	12.04	3.67	3.26	2.59	0.70
	Aspect (°)	80	0.00	357.51	174.36	135.00	130.72	0.75
MNP	LAI ($m^2.m^{-2}$)	68	0.47	5.00	1.90	1.90	0.84	0.44
	CCC ($\mu g.cm^{-2}$)	68	9.29	132.59	42.37	42.72	20.98	0.50
	FVC	68	0.25	0.97	0.62	0.60	0.18	0.28
	Grass height (cm)	68	4.50	37.00	16.38	15.75	8.30	0.51
	Elevation (m)	68	1307.59	1893.29	1470.54	1389.27	167.68	0.11
	Slope (°)	68	0.34	14.12	3.61	3.16	2.65	0.73
	Aspect (°)	68	0.00	354.81	131.63	68.11	125.11	0.95

Table 1 data for GGHNP (which is mainly a grassland environment) suggests the sampled subplots were characterised by high vegetation cover with little variability according to the respective mean, standard deviation and CV of the FVC. This is also corroborated by the high LAI values reaching a maximum of 6.24. The CCC suggests the grasses in the sampled areas were on average green and healthy. However, the grass height, LAI and CCC showed moderately high variability (i.e. according to CV that ranged from approximately 46% to 70%) compared to the FVC biophysical variable. This variability is representative of the various vegetation communities across the sections of the park, and in particular the heterogenous grassland environment. In addition, the CCC's variability could also be controlled by the different soils and climate within GGHNP, for example Li Ying et al. (2018).

On the other hand, the grasses in MNP (which is generally a mixed savanna and grassland environment) had on average, moderately-high vegetation cover marked by

FVC of 62%. The CCC was on average, moderate and showed little variability (StDev and CV of about 21% and 50% respectively) across the subplots in MNP compared to in GGHNP (Table 1). However, the LAI and grass height appeared to have a relatively high variability (i.e. CV of about 44% and 51% respectively) across the subplots indicating widespread vegetation structural differences and heterogeneity of the mountainous savanna and grassland environment.

Furthermore, Table 1 include statistical information on the altitude, slope and aspect particularly where our subplots were located. This provides important information on terrain variability where our subplots were located. It is evident that our study sites are mountainous, characterised by high altitude ranges (Table 1). Considerable care during sampling design was taken to ensure that our subplots were located in fairly homogeneous surroundings characterised by varying slope terrains as this can be seen on the slope values (Table 1). On average the two sites, at least where our subplots were located had south to southeast facing slopes i.e. aspect (Table 1). Overall, the field measured biophysical variables (Table 1) were used in this study to quantify the accuracy of the SL2P-derived LAI, CCC and FVC variables.

To further explore the variability of the field measurements, the grass canopy reflectance values from Sentinel-2 images were extracted from each subplot (Figure 6). The reflectance included the mean, minimum and maximum values in-order to show the spectral variability of the subplots in the two sites. The GGHNP spectral curves show that the subplots had high chlorophyll content based on the shape of the spectral profile in the visible, near-infrared and shortwave infrared regions respectively. The shape is more characteristic of greener vegetation than in MNP. This is consistent with the CCC presented in Table 1 for GGHNP which on average appeared higher and more variable than in MNP. Generally, the spectral variability may also have been influenced by the

various grass species within the subplots. Grass species identification within the subplots was carried out by the rangers in the two sites through visual interpretation, and this information was manually recorded and linked to the subplots.

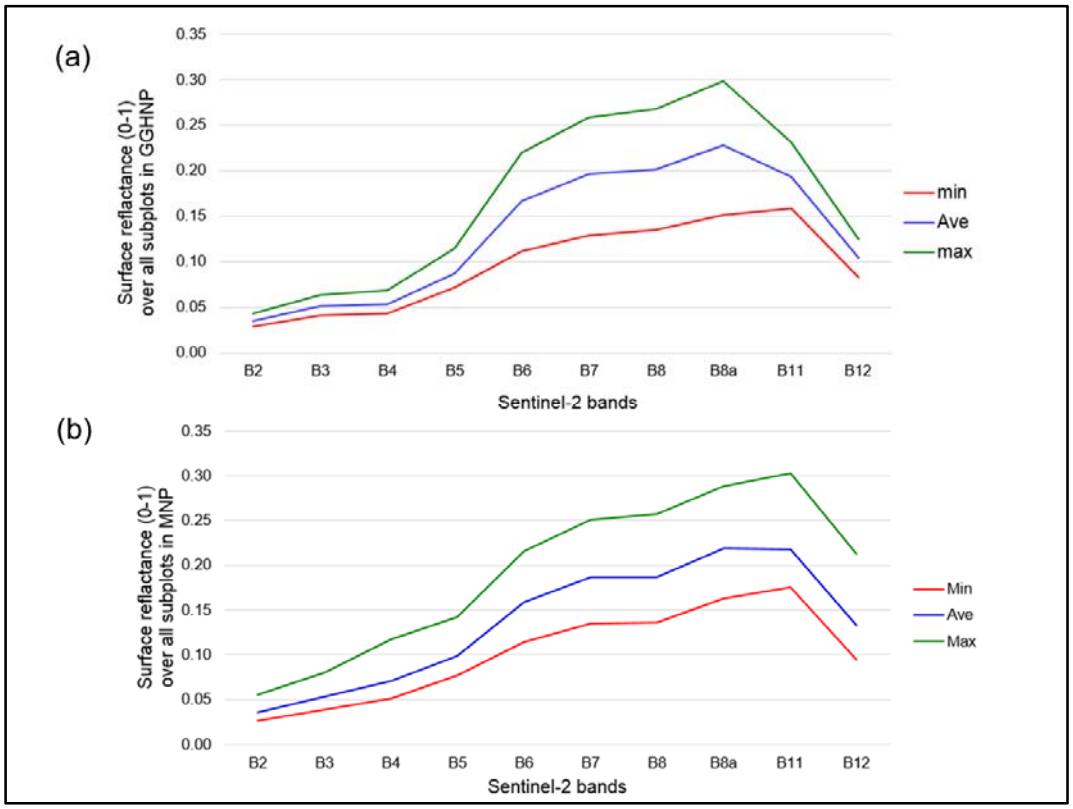

Figure 6. The Mean, mimimum and maximum (grass) canopy reflectance spectra showing the spectral variability of the subplots in the Golden Gate Highland National Park (a) and Marakele National Park (b) based on the Sentinel-2 bands B2: 458-523 nm, B3: 543-578 nm, B4: 650-680 nm, B5: 698-713 nm, B6: 733-748 nm, B7: 773-793 nm, B8: 785-899 nm, B8a: 855-875 nm, B11: 1565-1655 nm and B12: 2100-2280 nm.

Statistical analysis

The validation of the SL2P-derived LAI, CCC and FVC modelled estimates with field measurements was performed using the R^2, *RMSE*, *Bias* and mean absolute error (*MAE*). These represent some of the standard performance metrics (Chai and Draxler, 2014) used in numerous validation studies to quantify overall agreement and error between the data sets being compared, for example Brown et al. (2021), Hu et al. (2020), and Kganyago et al. (2020). In this study, a linear regression model was created for each pair of variables comprising the dependent and independent variables; and thereafter the

R^2 was computed for each model to measure the goodness of fit. This was followed by the computation of *RMSE* using Equation (2):

$$RMSE = \sqrt{\frac{\sum_{k=1}^{n} (e_k - m_k)^2}{n}} \quad (2)$$

where e_k represents the estimated value according to the SL2P, m_k represents the field measured value, and n denotes the number of observations available for analysis. *RMSE* can range from 0 to ∞ and the lower the value, the more accurate the SL2P-derived modelled estimates. To further assess the SL2P model performance, *MAE* was used an additional metric (Chai and Draxler, 2014) expressed in Equation (3):

$$MAE = \frac{1}{n} \sum_{k=1}^{n} | e_k - m_k | \quad (3)$$

The combination of *MAE* and *RMSE* metrics gave a representation of the variation in model error distribution, which can be normally- or uniformly distributed (Chai and Draxler, 2014). Another performance metric used for evaluating the SL2P biophysical models was the *Bias* calculated using Equation (4):

$$Bias = \frac{1}{n} \sum_{k=1}^{n} (e_k - m_k) \quad (4)$$

where a negative *Bias* implies a systematic underestimation of the field measurements by the SL2P biophysical modelled estimates (Brown et al., 2021). In contrast, a positive *Bias* implies a systematic overestimation. Additionally, the relative *Bias* (*pBias%*) and root mean square error (*RRMSE%*) were calculated by dividing each indicator (*Bias* or *RMSE*) with the mean of the field measurements.

Results

Accuracy assessment

Accuracy assessment in GGHNP

Figure 7 shows the accuracy of the SNAP SL2P biophysical estimates that were evaluated based on field measurements. The performance of the SL2P biophysical estimates differed with the field biophysical estimates in the GGHNP. In particular, poor agreement (low R^2 values) with field biophysical estimates was observed for the SL2P-derived LAI, FVC and CCC estimates. This was accompanied by the relatively high *RMSE* values of 1.83 $m^2.m^{-2}$ and 0.53 corresponding to the SL2P derived LAI and FVC estimates, respectively. However, *RMSE* of the SL2P CCC estimate was higher at about 39.23 $\mu g.cm^{-2}$ when evaluated based on the field measured CCC. Furthermore, the *RRMSE* of the SL2P LAI, FVC and CCC estimates was found to be 80.97%, 61.63% and 85.26% respectively. This suggest, the accuracy of the SL2P modelled estimates in GGHNP was considerably low. On the other hand, Figure 7 reveals some level of underestimations of the SL2P LAI (*Bias* = -1.31 $m^2.m^{-2}$), FVC (*Bias* = -0.51) and CCC (*Bias* = -14.8 $\mu g.cm^{-2}$) according to our field measurements. These corresponded to *pBias's* of about 57.96%, 59.30% and 32.17% respectively. Although, the underestimations were generally high, the SL2P CCC provided notably lower underestimation.

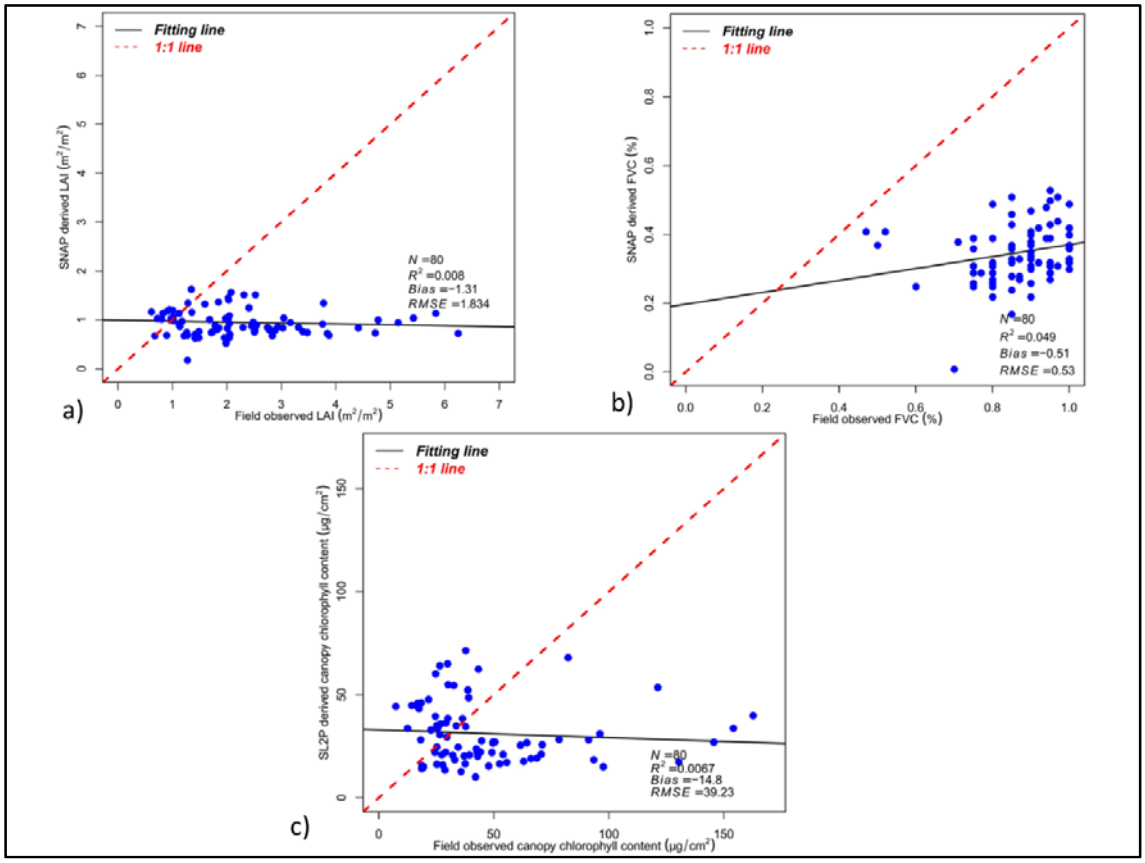

Figure 7. The biophysical-specific comparison between field measurements and the SNAP Sentinel-2 Level 2 Prototype Processor modelled estimates in GGHNP on 18 – 21 March 2021. The solid black line and red dotted line represent the linear fit for all extracted pixels and a 1:1 relationship, respectively for the LAI, FVC and CCC from (a) – (c).

Accuracy assessment in MNP

Figure 8 shows the performance of the SL2P biophysical estimates when compared with the field biophysical estimates in MNP. There is a relatively poor agreement (low R^2 values) between the SL2P modelled biophysical estimates and field measured estimates of LAI, FVC, CCC. In particular, the computed *RMSE* values of SL2P estimates indicated a general minor decline in error compared to the results in GGHNP (Figure 8). This decline is also evident in the computed *RRMSE* values of, especially SL2P LAI (78.95%) and CCC (62.36%) than FVC (67.74%). This is an indication of the low accuracy of SL2P modelled estimates in MNP. Additionally, the SL2P biophysical estimates had marginal underestimations of the field based

measurement of LAI (*Bias* = -1.20 $m^2.m^{-2}$), FVC (*Bias* = -0.37) and CCC (*Bias* = -15.6 $\mu g.cm^{-2}$), compared to in GGHNP. These *Bias values* corresponded to *pBias's* of about 63.16%, 59.68% and 36.91% respectively. Although, the underestimations were generally high, the SL2P CCC provided notably lower underestimation. Overall, the performance of the SL2P modelled estimates in the two study sites was characterised by low accuracy with *RRMSE* values ranging from 61.63% to 85.26% (Figure 7 and Figure 8) coupled with possible systematic underestimations with (negative) *pBias* values ranging from 32.17% to 63.16% (Figure 7 and Figure 8). Several studies on statistical evaluation of model accuracy categorised any RRMSE values greater than 30% as a reflection of poor model accuracy (Li et al., 2013, Bandyopadhyay et al., 2008, Jamieson et al., 1991). The same analogy can be extended to the RRMSE values revealed by the SL2P estimates in GGHNP and MNP.

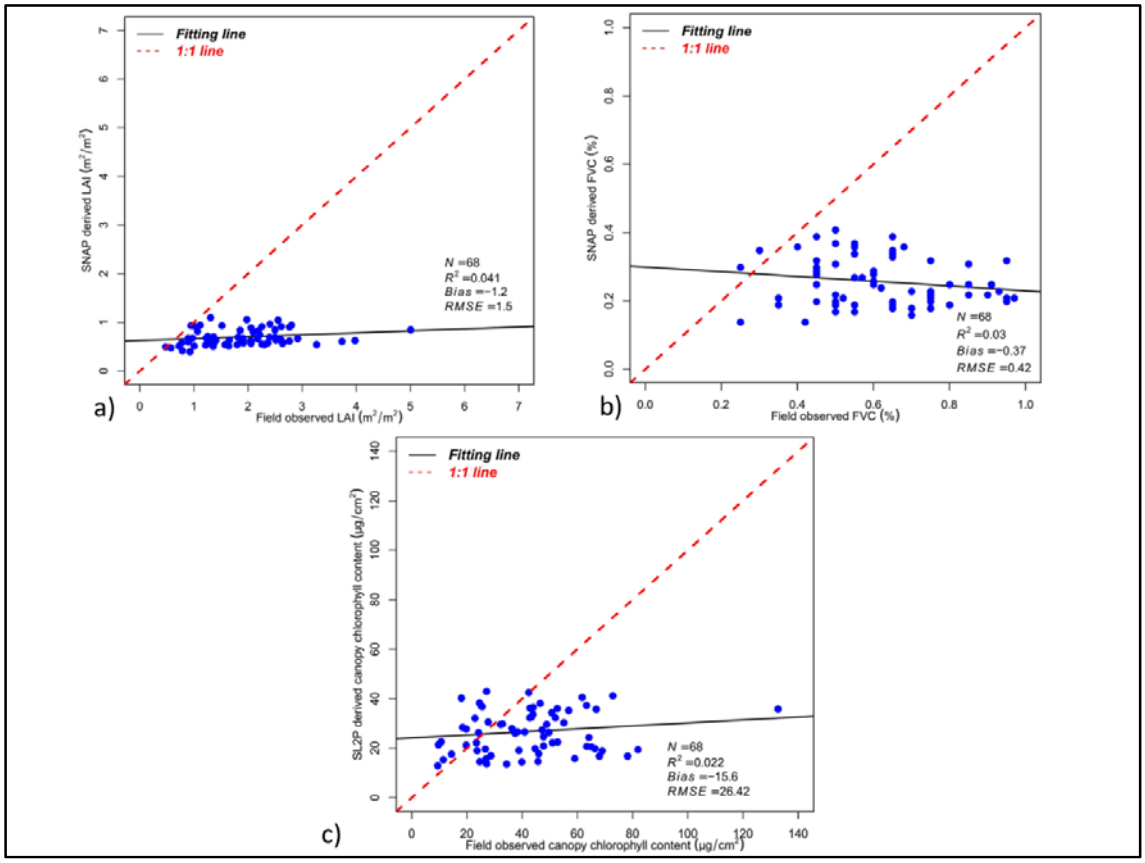

Figure 8. The biophysical-specific comparison between field measurements and the SNAP Sentinel-2 Level 2 Prototype Processor modelled estimates in MNP on 8 – 10 April 2021. The solid black line and red dotted line represent the linear fit for all extracted pixels and a 1:1 relationship, respectively for the LAI, FVC and CCC from (a) – (c).

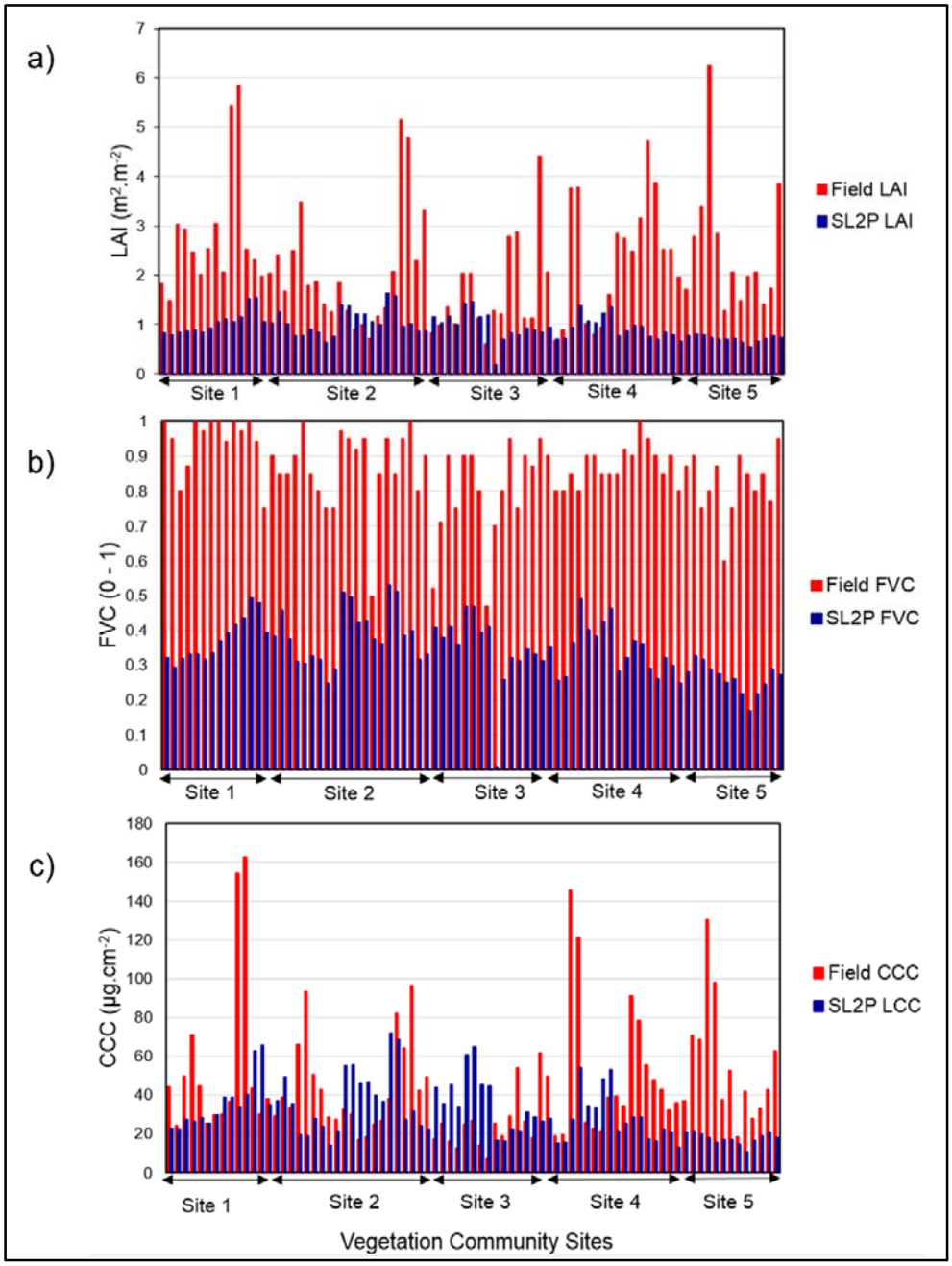

Figure 9. Histograms of field and SL2P biophysical estimates in GGHNP across the visited subplots in different sites.

Consistency of the three SL2P-derived biophysical estimates across the parks

Figure 9 shows the histograms of corresponding field and SL2P biophysical estimates in GGHNP across the visited subplots in different sites characterised by varying vegetation communities. The SL2P-derived estimates were generally consistent in

underestimating the LAI, FVC and CCC biophysical variables according to the field estimates in the visited sites of GGHNP. However, in some sites the SL2P CCC was not consistent in underestimating the field measured CCC. Across all visited sites, the SL2P-derived LAI and FVC had lower mean estimates of 0.95 $m^2.m^{-2}$ and 35% compared to field estimates of 2.26 $m^2.m^{-2}$ and 86% respectively. Additionally, the SL2P-derived LAI and FVC showed little variability across the sites with standard deviations of approximately 0.26 $m^2.m^{-2}$ and 8.7% when compared to field estimates i.e. standard deviations of 1.24 $m^2.m^{-2}$ and 11% respectively. A view of the selected pictures of the visited subplots across the sites show that, the areas in general had high grass coverage (Figure 10) which could also be seen in the high field estimates of LAI and FVC (Figure 9). This appears to have not been well represented or captured by the SL2P-derived estimates.

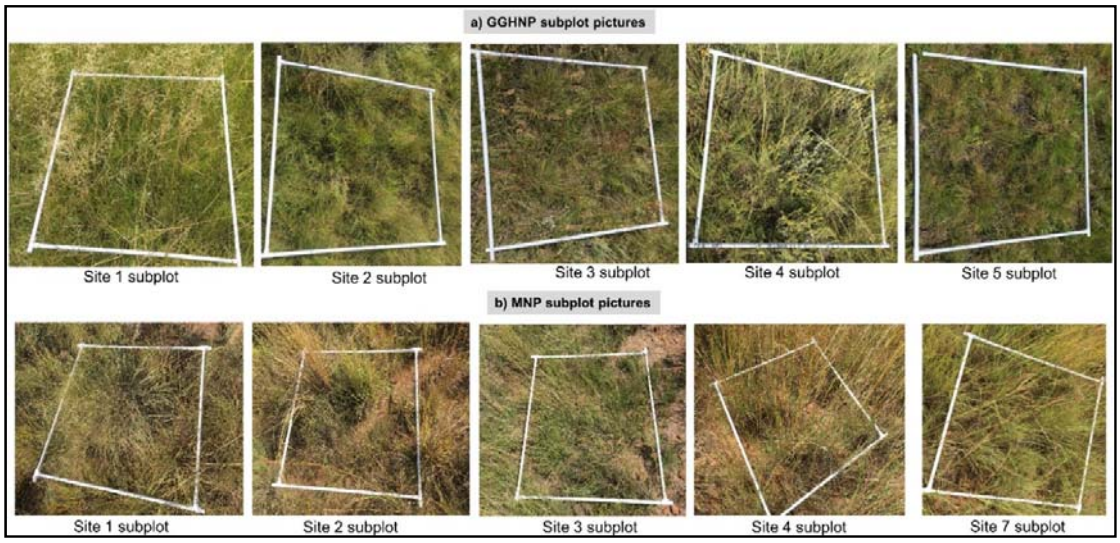

Figure 10. Selected sample of pictures of the visited subplots taken at Nadir across the sites in GGHNP (a) and MNP (a).

However, interestingly the varying shape distribution of the histograms suggests a trend of agreement between the fluctuating estimates of SL2P and field observed, in that, higher and lower SL2P-derived estimates followed the trend of the field estimates across all sites for CCC, FVC and LAI. Furthermore, the field-based CCC was overestimated, more especially in sites 2 and 3 (including a few subplots in sites 1 and 4) according to the SL2P-derived estimates of CCC. On average, the SL2P-derived CCC had a mean estimate of approximately 31.21 $\mu g.cm^{-2}$, which was lower compared to that of the field-based CCC at 46.01 $\mu g.cm^{-2}$. Furthermore, the SL2P-derived CCC showed low variability with a standard deviation of 14.77 $\mu g.cm^{-2}$ across the sites, compared to the field-based CCC with a standard deviation of 32.26 $\mu g.cm^{-2}$. This higher variability in the field measured CCC is probably reflective of the different vegetation species and their composition across the visited subplots (see Figure 10 as well as the spectral plots in Figure 6). This variability may not have been well represented by the SL2P. Notwithstanding the observed discrepancy between the estimates, the histogram (Figure 9) reveals a trend of agreement between the fluctuating estimates of SL2P and field observed CCC across the sites in GGHNP. Overall, this trend suggests that the SL2P model has the potential to capture the changes in the CCC, LAI and FVC present in the varying vegetation communities across the natural heterogenous landscape.

Figure 11 shows the histograms of corresponding field and SL2P biophysical estimates in MNP across the visited subplots in different sites characterised by varying vegetation communities. The SL2P-derived biophysical estimates were generally consistent in underestimating the LAI, FVC and CCC according to the field estimates in the visited sites of MNP. In all visited sites, the SL2P-derived LAI and FVC had lower mean estimates of 0.70 $m^2.m^{-2}$ and 25.60% compared to field-based estimates at 1.90 $m^2.m^{-2}$ and 62.24% respectively. Furthermore, the SL2P-derived LAI and FVC revealed

little variability across the sites (Figure 11) with standard deviations of approximately 0.16 $m^2.m^{-2}$ and 6.98% when compared to field estimates i.e. standard deviations of 0.84 $m^2.m^{-2}$ and 17.62% respectively. A view of the selected pictures of the visited subplots across the sites show that, MNP in general had moderately-high fractional cover of grasses that appeared taller (see grass height information in Table 1), with more portions of exposed background soil compared to in GGHNP (Figure 4). These observations can also be seen in the histograms of field-based FVC and LAI estimates (Figure 11).

In particular, the variability and magnitude of the field-based FVC and LAI estimates appears not to have been well represented by the SL2P-derived estimates. The histograms suggests a rather non-responsive trend of agreement between the fluctuating estimates of SL2P and field observed LAI and FVC. However, it is only in sites 4 and 5/6/7 where the SL2P- and field-based LAI and FVC fluctuations (histograms) revealed a subtle trend of agreement, despite the discrepancy in the magnitude of the estimates. On the other hand, the field-based CCC was generally underestimated according to the SL2P-derived estimates of CCC in all the sites. The histograms (Figure 11) showed a clear trend of agreement coupled with intermittent overlaps between the fluctuating estimates of SL2P and field-observed CCC in MNP across all the sites. This trend suggests, the SL2P in MNP responded fairly well by capturing the variability of the CCC field estimates across the sites.

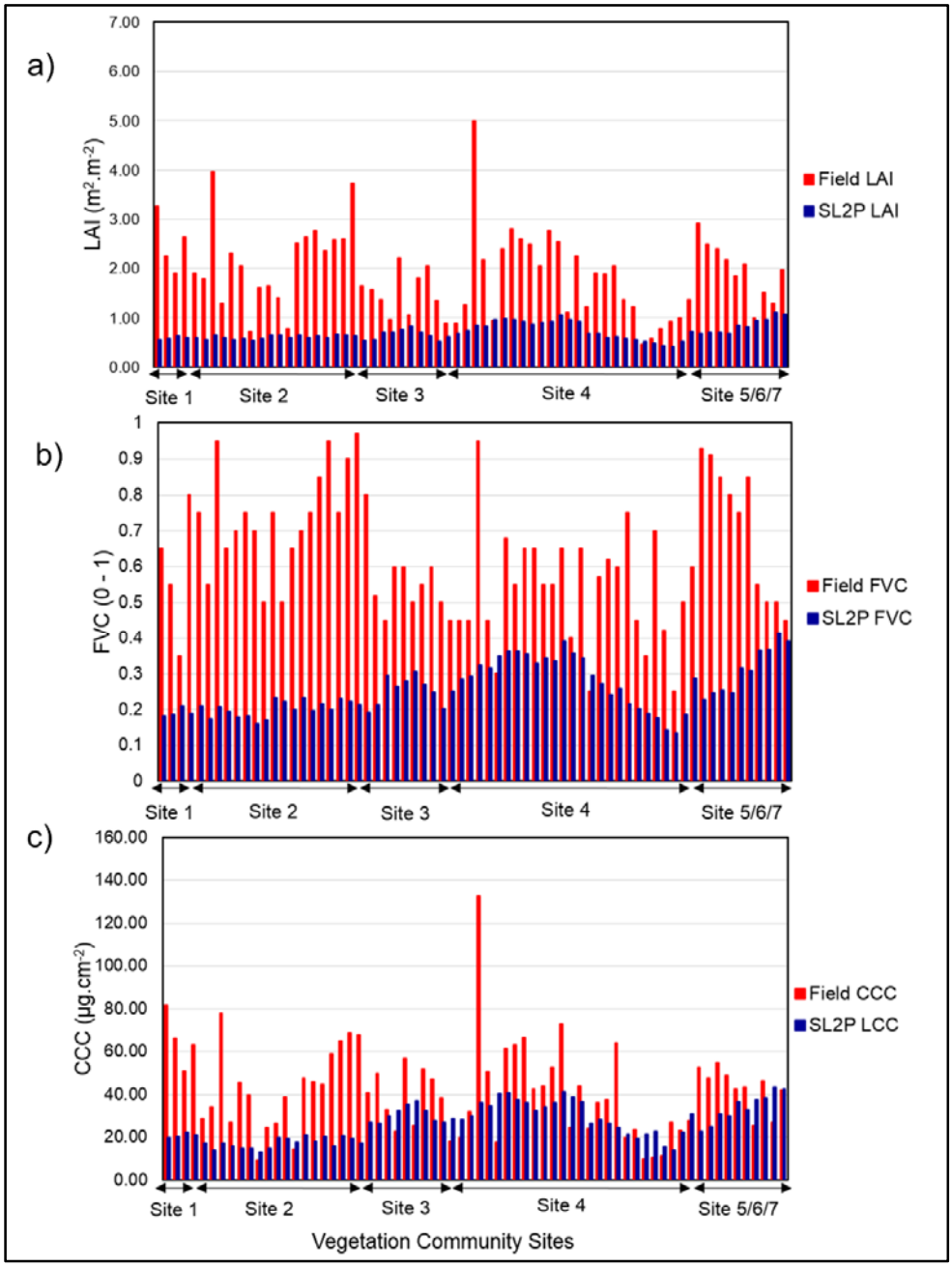

Figure 11. Histograms of field and SL2P biophysical estimates in MNP across the visited subplots in different sites.

On average, the SL2P-derived CCC had a mean estimate of approximately 26.74 µg.cm^{-2}, which was lower compared to that of the field-based CCC at 42.37 µg.cm^{-2}. The SL2P-derived CCC showed lower variability with a standard deviation of 8.56 µg.cm^{-2} across the sites, compared to the field-based CCC with a standard deviation of 20.98

$\mu g.cm^{-2}$ (Figure 11). This variability could be attributed to the variety of vegetation communities (Figure 1 and Figure 4), grass species (Figure 6) and terrain changes (Table 1) across the visited subplots.

Discussion

The results of the performance and consistency of the SL2P-derived LAI, FVC and CCC biophysical estimates over the heterogenous grassland canopy and landscape of two national parks in South Africa i.e. GGHNP and MNP, indicate that these estimates, derived at a spatial resolution of 20 m had moderate-to-high errors and appeared linearly biased. In particular, we explore a variety of potential underlying reasons that caused the discrepancies between the SL2P- and field-based biophysical estimates, in an attempt to better comprehend the performances of Sentinel-2 biophysical estimates.

Firstly, the Sen2Cor atmospheric correction procedure used in this study to pre-process the GGHNP and MNP Sentinel-2 data, reportedly has residual uncertainty for surface reflectance products in the approximate range from 0.012 to 0.017 which corresponds to the spectral domain of the following eight bands B3, B4, B5, B6, B7, B8a, B11 and B12 (Hu et al., 2020, Djamai and Fernandes, 2018, Li Yingjie et al., 2018, Martins et al., 2017). These bands are used by the SNAP SL2P for generation of the 20 m resolution biophysical estimates (Weiss and Baret, 2020) and therefore the magnitude of uncertainties or errors in the SL2P biophysical estimates of LAI, FVC and CCC found in this study, can to some extent be attributed to the residual error of the atmospheric correction procedure.

Secondly, some of the errors in the SL2P estimates found in this study, could be attributed to known uncertainties associated with the SL2P adopted ANNs that are trained with RTM simulations. According to Weiss and Baret (2020) the SL2P algorithm provide estimations of LAI, FVC and CCC with RMSEs of 0.90 $m^2.m^{-2}$, 0.041, 57.99 $\mu g.cm^{-2}$

respectively. Perhaps this could be used to explain the generally high uncertainties found in the SL2P-derived LAI and CCC estimates, while the FVC errors (*RMSE* and *RRMSE*) in the two parks were the lowest coupled with *Bias* values that were closest to zero. For example, based on the SL2P algorithm evaluation by Weiss and Baret (2020), the errors of LAI and CCC estimates were found to escalate with increasing LAI and CCC values. A similar effect was observed in our findings whereby the SL2P-derived LAI and CCC estimates in GGHNP had higher errors (i.e. RMSEs of 1.834 $m^2.m^{-2}$ and 39.23 $\mu g.cm^{-2}$) than those found in MNP with RMSEs of 1.50 $m^2.m^{-2}$ and 26.42 $\mu g.cm^{-2}$ respectively (Figure 7 and Figure 8). In particular, our field measurements of LAI and CCC in GGHNP revealed the highest maximum values of 6.24 $m^2.m^{-2}$ and 162.61$\mu g.cm^{-2}$ compared to the maximum values in MNP of 5.00 $m^2.m^{-2}$ and 132.59 $\mu g.cm^{-2}$, respectively (Table 1). A study by Brown et al. (2021) found that the SL2P-derived LAI appears to perform comparatively poorly, particularly over heterogenous canopies with LAI values greater than 3. Furthermore, Hu et al. (2020) found that the SL2P-derived LAI and FVC estimates tend to systematically underestimate grasses (i.e. giving negative *Bias* values), and suggested increasing the number of ground measurements for a comprehensive evaluation of the SL2P biophysical products. This observation on possible systematic underestimation corroborates the findings of this study. Similarly, the SL2P-derived CCC underestimated field-based CCC in both GGHNP and MNP with *Bias* values of -14.80 $\mu g.cm^{-2}$ and -15.64 $\mu g.cm^{-2}$ (Figure 7 and Figure 8). It remains an exercise for further investigation on whether this underestimation is systematic, particularly over heterogeneous grassland canopies and landscapes (Ali et al., 2020). Furthermore, it is worth mentioning that the empirical calibration method described in Markwell et al. (1995) is for soybean and maize leaves, so there could be some species-specific biases when applying to other grass species. An interesting observation was that the variability

of the SL2P CCC estimates, appeared to be representative of the grass species heterogeneity over the visited sites, more especially in GGHNP (Figure 9) where about 36 different grass species were detected within the subplots. For example, in the 36 grass species identified within the subplots in GGHNP, the most dominant were found to be, in chronological order: *Eragrostis curvula*, *Elionurus muticus*, *Aristida adscensionis*, *Stiburus alopecuroide*, *Sporobolus africanus*, *Heteropogon contortus*, *Tristachya leucothrix*, *Microchloa caffra*, *Themeda triandra*, *Urochloa decumbens*, *Helichrysum rugulosum* and *Helichrysum pilosellum*.

Thirdly, the temporal mismatch between the Sentinel-2 data acquisition dates and the field data collection dates is also another factor, which may have contributed to the SL2P biophysical estimates presented in this study. Although considerable efforts were made in this study to minimize as much as possible the temporal gaps between the data sets.

Lastly, since the SL2P biophysical estimates of LAI, FVC and CCC had the lower errors in MNP compared to in GGHNP according to the field-based measurements; it may be premature to conclude at this stage whether the SL2P biophysical estimates are more reliable in MNP. More field-based measurements ought to be obtained in future to confirm the consistency and reliability of the SL2P estimates in the two parks. For example, 68 and 80 samples were obtained in MNP and GGHNP respectively, and therefore more samples (field measurements of biophysical estimates) should be acquired not only in MNP where fewer samples were obtained in this study, but also in both parks in-order to ameliorate the robustness of the validation results. Recent validation studies showed that the performance of SL2P biophysical estimates generally satisfy user requirements in various vegetation types that are characterised by homogenous canopies (Brown et al., 2021). However, the performance of the SL2P biophysical estimates in the

grasses located in heterogeneous natural environments is yet to be fully understood; hence the need for more multi-date and well-distributed ground measurements of biophysical estimates in such environments (Cho et al., 2014, Hu et al., 2020).

Conclusion

The aim of this study was to validate the SL2P retrieved LAI, FVC and CCC estimates over a heterogenous savanna and grassland environment for the wet season of 2021 in South Africa using field measurements. Overall, the results show that the current high-resolution Sentinel-2 derived biophysical estimates demonstrated comparatively poor performance over the grasslands in both study sites i.e. MNP and GGHNP. The SL2P biophysical estimates of LAI, FVC and CCC had the lowest errors in MNP compared to in GGHNP according to the field-based measurements. Furthermore, the SL2P LAI, FVC and CCC estimates appeared to systematically underestimate the grasses (i.e. indicated by negative *Bias* values) in both study sites according to our field estimates. Nonetheless, we noted that the variability of the SL2P-derived CCC estimates, appeared to be responsive of the grass species heterogeneity over the visited sites. The performance of the SL2P-derived biophysical estimates in both study sites was consistent in that the estimates appeared to be linearly biased, coupled with moderate-to-high errors. Other existing validation studies of the SL2P estimates made virtually similar remarks regarding their performance over grasses, especially in a heterogenous canopy for example, Brown et al. (2021), and Hu et al. (2020).

Several underlying reasons that could have potentially contributed the discrepancies between the SL2P- and field-based biophysical estimates were identified as follows: (i) the residual uncertainties emanating from the Sen2Cor atmospheric correction procedure, (ii) the inherent uncertainties associated with the SL2P adopted

ANNs that are trained with RTM simulations, (iii) the temporal mismatch between the Sentinel-2 data acquisition dates and the field data collection dates, and (iv) limited samples in-terms of quantity and spatiotemporal distribution, and (v) field data collection procedures especially with regards to using the ACCUPAR for LAI_e estimation, which partially corrects the effect of foliage clumping and could have possibly overestimated the actual green LAI. The reasons above suggests the need for extensive validation of the SL2P LAI, FVC and CCC estimates over natural heterogenous canopies using long-term field measurements in different regions. In addition, there may also be a need for improving the current SL2P algorithm and/ or investigating other inversion methods to provide better estimates with reduced biases and error over the aforementioned canopies.

Author Contributions: Conceptualization, P.T. and A.R.; methodology, P.T. and A.R.; Formal analysis, P.T. and M.Q; validation, P.T., M.Q., and M.M.; resources, G.C.; writing—original draft preparation, P.T.; writing—review and editing, A.R., M.Q., M.M., G.C; project administration, P.T.;

Funding: This research was funded by Research development programme of the University of Pretoria, as well as the Southern African Science Service Centre for Climate Change and Adaptive Land Management (SASSCAL) of the National Research Foundation (NRF) of South Africa, grant number 118590.

Acknowledgments: The Sentinel-2 data used in this study were downloaded from the European Space Agency Copernicus Open Access Hub. We sincerely thank the field assistants (namely Phomolo Seriba, Katlego Mashiane, Steven Khosa and Brian Mabunda) in the Golden Gate Highlands National Park and Marakele National Park for their collaborative effort in collecting grass LAI, FVC and LCC ground measurements. Furthermore, we sincerely thank the Agricultural Research Council of

South Africa for their assistance with filed data collection equipment. Last but not least, we thank the anonymous reviewers for their helpful comments on the initial version of this manuscript.

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

Data availability statement: We understand that the publication of the data is becoming a good practice in research. However, we plan to share all our data in future, but at this stage we are still going to further analyse it for locally parameterized types of models, looking at both empirical and the inversion of the physically-based models.

References

Intergovernmental Science-Policy Platform on Biodiversity and Ecosystem Services (IPBES). Summary for Policymakers of the Global Assessment Report on Biodiversity and Ecosystem Services of the Intergovernmental Science-Policy Platform on Biodiversity and Ecosystem Services (IPBES Secretariat, Bonn), 2019, Available online: https://ipbes.net/sites/default/files/2020-02/ipbes_global_assessment_report_summary_for_policymakers_en.pdf (accessed on 21 July 2021).

Global Biodiversity Outlook 5 (GBO05). CBD (Convention on Biological Diversity). CBD Secretariat, Montréal, Canada, 2020, Available online: https://www.cbd.int/gbo/gbo5/publication/gbo-5-en.pdf (accessed on 20 August 2021).

Living Planet Report 2020 - Bending the curve of biodiversity loss, 2020, Available online: https://f.hubspotusercontent20.net/hubfs/4783129/LPR/PDFs/ENGLISH-FULL.pdf (accessed on 29 September 2021).

Climate Change 2021: The Physical Science Basis. Contribution of Working Group I to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change, 2021,

Available online: https://www.ipcc.ch/report/ar6/wg1/downloads/report/IPCC_AR6_WGI_Full_Report_smaller.pdf (accessed on 29 September 2021).

Global Climate Observing System (GCOS). Systematic Observation Requirements for Satellite-Based Products for Climate, 2011 Update, Supplemental Details to the Satellite-Based Component of the Implementation Plan for the Global Observing System for Climate in Support of the UNFCCC (2010 Update). Available online: https://library.wmo.int/doc_num.php?explnum_id=3710 (accessed on 28 September 2021).

Ali AM, Darvishzadeh R, Skidmore A, Gara TW, O'Connor B, Roeoesli C, Heurich M, Paganini M. 2020. Comparing methods for mapping canopy chlorophyll content in a mixed mountain forest using Sentinel-2 data. International Journal of Applied Earth Observation and Geoinformation. 87:102037.

Ali AM, Darvishzadeh R, Skidmore A, Heurich M, Paganini M, Heiden U, Mücher S. 2020. Evaluating prediction models for mapping canopy chlorophyll content across biomes. Remote sensing. 12(11):1788.

Bacour C, Baret F, Béal D, Weiss M, Pavageau K. 2006. Neural network estimation of LAI, fAPAR, fCover and LAI× Cab, from top of canopy MERIS reflectance data: Principles and validation. Remote sensing of environment. 105(4):313-325.

Bandyopadhyay A, Bhadra A, Raghuwanshi N, Singh R. 2008. Estimation of monthly solar radiation from measured air temperature extremes. Agricultural and forest meteorology. 148(11):1707-1718.

Baret F, Weiss M, Lacaze R, Camacho F, Makhmara H, Pacholcyzk P, Smets B. 2013. GEOV1: LAI and FAPAR essential climate variables and FCOVER global time series

capitalizing over existing products. Part1: Principles of development and production. Remote sensing of environment. 137:299-309.

Bei C, ZHAO Q-j, HUANG W-j, SONG X-y, YE H-c, ZHOU X-f. 2019. Leaf chlorophyll content retrieval of wheat by simulated RapidEye, Sentinel-2 and EnMAP data. Journal of Integrative Agriculture. 18(6):1230-1245.

Brown LA, Fernandes R, Djamai N, Meier C, Gobron N, Morris H, Canisius F, Bai G, Lerebourg C, Lanconelli C. 2021. Validation of baseline and modified Sentinel-2 Level 2 Prototype Processor leaf area index retrievals over the United States. ISPRS Journal of Photogrammetry and Remote Sensing. 175:71-87.

Brown LA, Ogutu BO, Dash J. 2019. Estimating forest leaf area index and canopy chlorophyll content with Sentinel-2: An evaluation of two hybrid retrieval algorithms. Remote Sensing. 11(15):1752.

Campbell GS. 1986. Extinction coefficients for radiation in plant canopies calculated using an ellipsoidal inclination angle distribution. Agricultural and forest meteorology. 36(4):317-321.

mCG. 1997. 1: 1 000 000 Scale Geological Map of the Republic of South Africa and the Kingdoms of Lesotho and Swaziland. Council for Geosciences Pretoria.

Chai T, Draxler RR. 2014. Root mean square error (RMSE) or mean absolute error (MAE)?–Arguments against avoiding RMSE in the literature. Geoscientific model development. 7(3):1247-1250.

Chen J, Menges C, Leblanc S. 2005. Global mapping of foliage clumping index using multi-angular satellite data. Remote Sensing of Environment. 97(4):447-457.

Estimation of leaf area index (LAI) of South Africa from MODIS imager by inversion of PROSAIL radiative transfer model. 2014 IEEE Geoscience and Remote Sensing Symposium; 2014: IEEE.

Confalonieri R, Francone C, Foi M. 2014. The PocketLAI smartphone app: an alternative method for leaf area index estimation.

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 Sensing of Environment. 112(5):2592-2604.

Darvishzadeh R, Wang T, Skidmore A, Vrieling A, O'Connor B, Gara TW, Ens BJ, Paganini M. 2019. Analysis of Sentinel-2 and RapidEye for Retrieval of Leaf Area Index in a Saltmarsh Using a Radiative Transfer Model. Remote sensing. 11(6):671.

Dash J, Curran P. 2004. The MERIS terrestrial chlorophyll index.

Disney M, Muller J-P, Kharbouche S, Kaminski T, Voßbeck M, Lewis P, Pinty B. 2016. A new global fAPAR and LAI dataset derived from optimal albedo estimates: Comparison with MODIS products. Remote Sensing. 8(4):275.

Djamai N, Fernandes R. 2018. Comparison of SNAP-derived Sentinel-2A L2A product to ESA product over Europe. Remote Sensing. 10(6):926.

Djamai N, Fernandes R, Weiss M, McNairn H, Goïta K. 2019. Validation of the Sentinel Simplified Level 2 Product Prototype Processor (SL2P) for mapping cropland biophysical variables using Sentinel-2/MSI and Landsat-8/OLI data. Remote sensing of environment. 225:416-430.

Fang H, Li W, Wei S, Jiang C. 2014. Seasonal variation of leaf area index (LAI) over paddy rice fields in NE China: Intercomparison of destructive sampling, LAI-2200, digital hemispherical photography (DHP), and AccuPAR methods. Agricultural and Forest Meteorology. 198:126-141.

Feret J-B, François C, Asner GP, Gitelson AA, Martin RE, Bidel LP, Ustin SL, Le Maire G, Jacquemoud S. 2008. PROSPECT-4 and 5: Advances in the leaf optical properties

model separating photosynthetic pigments. Remote sensing of environment. 112(6):3030-3043.

mFilipponi F. 2021. Comparison LAI estimates from of high resolution satellite observations using different biophysical processors. Pro-ceedings.

Frampton WJ, Dash J, Watmough G, Milton EJ. 2013. Evaluating the capabilities of Sentinel-2 for quantitative estimation of biophysical variables in vegetation. ISPRS journal of photogrammetry and remote sensing. 82:83-92.

Francone C, Pagani V, Foi M, Cappelli G, Confalonieri R. 2014. Comparison of leaf area index estimates by ceptometer and PocketLAI smart app in canopies with different structures. Field Crops Research. 155:38-41.

García-Haro FJ, Campos-Taberner M, Munoz-Mari J, Laparra V, Camacho F, Sanchez-Zapero J, Camps-Valls G. 2018. Derivation of global vegetation biophysical parameters from EUMETSAT Polar System. ISPRS journal of photogrammetry and remote sensing. 139:57-74.

Gastwirth JL. 1971. On the sign test for symmetry. Journal of the American Statistical Association. 66(336):821-823.

Gitelson AA, Vina A, Ciganda V, Rundquist DC, Arkebauer TJ. 2005. Remote estimation of canopy chlorophyll content in crops. Geophysical Research Letters. 32(8).

Gitelson AA, Viña A, Verma SB, Rundquist DC, Arkebauer TJ, Keydan G, Leavitt B, Ciganda V, Burba GG, Suyker AE. 2006. Relationship between gross primary production and chlorophyll content in crops: Implications for the synoptic monitoring of vegetation productivity. Journal of Geophysical Research: Atmospheres. 111(D8).

Hauser LT, Féret J-B, Binh NA, van der Windt N, Sil ÂF, Timmermans J, Soudzilovskaia NA, van Bodegom PM. 2021. Towards scalable estimation of plant functional diversity

from Sentinel-2: In-situ validation in a heterogeneous (semi-) natural landscape. Remote Sensing of Environment. 262:112505.

Hill MJ, Senarath U, Lee A, Zeppel M, Nightingale JM, Williams RDJ, McVicar TR. 2006. Assessment of the MODIS LAI product for Australian ecosystems. Remote Sensing of Environment. 101(4):495-518.

Hu Q, Yang J, Xu B, Huang J, Memon MS, Yin G, Zeng Y, Zhao J, Liu K. 2020. Evaluation of global decametric-resolution LAI, FAPAR and FVC estimates derived from Sentinel-2 imagery. Remote Sensing. 12(6):912.

Jamieson P, Porter J, Wilson D. 1991. A test of the computer simulation model ARCWHEAT1 on wheat crops grown in New Zealand. Field crops research. 27(4):337-350.

Jia K, Yang L, Liang S, Xiao Z, Zhao X, Yao Y, Zhang X, Jiang B, Liu D. 2018. Long-term Global Land Surface Satellite (GLASS) fractional vegetation cover product derived from MODIS and AVHRR Data. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing. 12(2):508-518.

Johnson MVV, Kiniry JR, Burson BL. 2010. Ceptometer deployment method affects measurement of fraction of intercepted photosynthetically active radiation. Agronomy journal. 102(4):1132-1137.

Jonckheere I, Fleck S, Nackaerts K, Muys B, Coppin P, Weiss M, Baret F. 2004. Review of methods for in situ leaf area index determination: Part I. Theories, sensors and hemispherical photography. Agricultural and forest meteorology. 121(1-2):19-35.

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(10):883-892.

Lang A, Yueqin X. 1986. Estimation of leaf area index from transmission of direct sunlight in discontinuous canopies. Agricultural and forest Meteorology. 37(3):229-243

Li M-F, Tang X-P, Wu W, Liu H-B. 2013. General models for estimating daily global solar radiation for different solar radiation zones in mainland China. Energy conversion and management. 70:139-148.

Li Y, Chen J, Ma Q, Zhang HK, Liu J. 2018. Evaluation of Sentinel-2A surface reflectance derived using Sen2Cor in North America. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing. 11(6):1997-2021.

Li Y, He N, Hou J, Xu L, Liu C, Zhang J, Wang Q, Zhang X, Wu X. 2018. Factors influencing leaf chlorophyll content in natural forests at the biome scale. Frontiers in Ecology and Evolution. 6:64.

Liang S, Wang J. 2019. Advanced remote sensing: terrestrial information extraction and applications. Academic Press.

Sentinel-2 Sen2Cor: L2A processor for users. Proceedings Living Planet Symposium 2016; 2016: Spacebooks Online.

Lv T, Zhou X, Tao Z, Sun X, Wang J, Li R, Xie F. 2021. Remote Sensing-Guided Spatial Sampling Strategy over Heterogeneous Surface Ground for Validation of Vegetation Indices Products with Medium and High Spatial Resolution. Remote Sensing. 13(14):2674.

Markwell J, Osterman JC, Mitchell JL. 1995. Calibration of the Minolta SPAD-502 leaf chlorophyll meter. Photosynthesis research. 46(3):467-472.

Martins VS, Barbosa CCF, De Carvalho LAS, Jorge DSF, Lobo FdL, Novo EMLdM. 2017. Assessment of atmospheric correction methods for Sentinel-2 MSI images applied to Amazon floodplain lakes. Remote Sensing. 9(4):322.

Masemola C, Cho MA, Ramoelo A. 2016. Comparison of Landsat 8 OLI and Landsat 7 ETM+ for estimating grassland LAI using model inversion and spectral indices: case study of Mpumalanga, South Africa. International Journal of Remote Sensing. 37(18):4401-4419.

Mucina L, Rutherford MC. 2006. The vegetation of South Africa, Lesotho and Swaziland. South African National Biodiversity Institute.

Myneni RB, Hoffman S, Knyazikhin Y, Privette J, Glassy J, Tian Y, Wang Y, Song X, Zhang Y, Smith G. 2002. Global products of vegetation leaf area and fraction absorbed PAR from year one of MODIS data. Remote sensing of environment. 83(1-2):214-231.

Pastor-Guzman J, Brown L, Morris H, Bourg L, Goryl P, Dransfeld S, Dash J. 2020. The sentinel-3 OLCI terrestrial chlorophyll index (OTCI): Algorithm improvements, spatiotemporal consistency and continuity with the MERIS archive. Remote Sensing. 12(16):2652.

Ramoelo A, Cho MA. 2018. Explaining leaf nitrogen distribution in a semi-arid environment predicted on Sentinel-2 imagery using a field spectroscopy derived model. Remote Sensing. 10(2):269.

Ramoelo A, Cho MA, 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. International journal of applied earth observation and geoinformation. 43:43-54.

Schloderer G, Bingham M, Awange JL, Fleming KM. 2011. Application of GNSS-RTK derived topographical maps for rapid environmental monitoring: a case study of Jack Finnery Lake (Perth, Australia). Environmental monitoring and assessment. 180(1):147-161.

Skidmore AK, Coops NC, Neinavaz E, Ali A, Schaepman ME, Paganini M, Kissling WD, Vihervaara P, Darvishzadeh R, Feilhauer H. 2021. Priority list of biodiversity metrics to observe from space. Nature ecology & evolution. 5(7):896-906.

Svinurai W, Hassen A, Tesfamariam E, Ramoelo A. 2021. Modelled effects of grazing strategies on native grass production, animal intake and growth in Brahman steers. African Journal of Range & Forage Science.1-11.

Van Engelen V, Dijkshoorn J. 2013. Global and national soils and terrain digital databases (SOTER). Report-ISRIC World Soil Information.(2013/04).

Van Staden P, Bredenkamp G. 2005. Major plant communities of the Marakele National Park. Koedoe. 48(2):59-70.

Verhoef W. 1984. Light scattering by leaf layers with application to canopy reflectance modeling: the SAIL model. Remote sensing of environment. 16(2):125-141.

Wang X, Jia K, Liang S, Li Q, Wei X, Yao Y, Zhang X, Tu Y. 2017. Estimating fractional vegetation cover from landsat-7 ETM+ reflectance data based on a coupled radiative transfer and crop growth model. IEEE Transactions on Geoscience and Remote Sensing. 55(10):5539-5546.

Weiss M, Baret F. S2ToolBox Level 2 products: LAI, FAPAR, FCOVER. Version 2.0. 2020. Available online: 762 http://step.esa.int/docs/extra/ATBD_S2ToolBox_V2.0.pdf (accessed on 15 July 2021)

Xie Q, Dash J, Huete A, Jiang A, Yin G, Ding Y, Peng D, Hall CC, Brown L, Shi Y. 2019. Retrieval of crop biophysical parameters from Sentinel-2 remote sensing imagery. International Journal of Applied Earth Observation and Geoinformation. 80:187-195.