Improved retrieval of land surface biophysical variables from time series of Sentinel-3 OLCI TOA spectral observations by considering the temporal autocorrelation of surface and atmospheric properties

Peiqi Yang *, Wout Verhoef, Egor Prikaziuk, Christiaan van der Tol

Faculty of Geo-Information Science and Earth Observation (ITC), University of Twente, Enschede 7500 AE, the Netherlands

ARTICLE INFO

Keywords:
Temporal autocorrelation
Time series
Sentinel-3
OLCI
LAI
Radiative transfer model
SPART
Model inversion

ABSTRACT

Estimation of essential vegetation properties from remote sensing is crucial for a quantitative understanding of the Earth system. Ill-posedness of the model inversion problem leads to multiple interpretations of one satellite observation, and using prior information is a promising way to reduce the ill-posedness and increase the accuracy of land surface products. Tobler’s first law of geography states that “everything is related to everything else, but near things are more related than distant things”. Likewise, it is expected that the state of an object at a single moment is related to the state at every other moment, but temporally near attributes are more related than distant ones. This temporal autocorrelation is a vital source of prior information and can be used to improve the retrieval accuracy. In this study, we develop a retrieval framework that makes use of the temporal autocorrelation and dependence of land surface and atmospheric properties. We apply this retrieval algorithm to Sentinel-3 Ocean and Land Colour Instrument (OLCI) satellite data to derive land surface biophysical variables with a focus on leaf area index (LAI) from top-of-atmosphere (TOA) radiance observations. The results from both a synthetic dataset and a real satellite dataset show that the use of the temporal continuity as a priori information improves the accuracy of the estimation of land surface properties, such as leaf chlorophyll content and LAI. Compared with the MODIS LAI products, much less unrealistic short-term fluctuations are found in the LAI retrievals from OLCI with the time-series retrieval approach across different land cover types including cropland, forest and savannah. Field measurements of LAI at two forest sites quantitatively confirm that the estimated LAI from OLCI is reasonably accurate with $R^2 > 0.65$ and RMSE $< 1.00$ $m^2/m^2$. Overall, the time series retrieval results in more robust and smoother time series than standard retrievals of LAI from individual scenes, more stable retrievals than the MODIS LAI product, and values of LAI that match better with reported measurements in the field. The present retrieval framework can make better use of time series of spectral observations and potentially of multi-sensor observations.

1. Introduction

Vegetation is a dynamic component in the Earth system and an essential factor in land-atmosphere interactions. Characterizing key vegetation properties, such as leaf area index (LAI) and leaf chlorophyll content ($C_{ab}$), is essential to understand the vital role of vegetation in the Earth’s biosphere, because these properties largely determine the fraction of absorbed photosynthetically active radiation (fAPAR) and land surface energy budget (Sellers et al. 1997). Optical remote sensing provides a great opportunity for vegetation monitoring by means of measuring electromagnetic radiation reflected and emitted from vegetation (Baret and Buis 2008; Verrelst et al. 2019; Wang et al. 2019). Translation of remote sensing signals to vegetation properties are often based on knowledge gained from radiative transfer modelling or empirical relationships between vegetation properties and remote sensing signals. Correspondingly, physically-based (e.g., look-up tables and numerical optimization) and statistical (e.g., regression models based on vegetation indices) approaches have been developed to retrieve vegetation biophysical variables (Berger et al. 2020; Combal et al. 2003; Darvishzadeh et al. 2008; De Grave et al. 2020; Verrelst et al. 2015; Xiao et al. 2014). Compared with statistical approaches, physically-based approaches have clear advantages for global

* Corresponding author.
E-mail address: p.yang@utwente.nl (P. Yang).

https://doi.org/10.1016/j.rse.2021.112328
Received 9 October 2020; Received in revised form 29 January 2021; Accepted 31 January 2021
Available online 11 February 2021
0034-4257/© 2021 The Author(s). Published by Elsevier Inc. This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/).
applications, because they are based on physical laws and therefore generally applicable. Physically-based approaches rely on radiative transfer models (RTM) to link vegetation properties with remote sensing signals for various sun-observer geometries. Inversion of RTMs allows translating remote sensing signals to vegetation biophysical variables, but these models are usually nonlinear and sometimes complex. Simulated look-up tables (LUT) and trained artificial neural networks (ANN) are common tools to simplify and replace a complex RTM to improve the efficiency of inversion, but some accuracy can be lost due to the simplification. A more challenging and fundamental issue encountered in practice is the so-called ill-posedness of model inversion problems (Combol et al. 2003; Quaife and Lewis 2010; Tarantola 2005; Weiss et al. 2000), which is commonly expressed by the possibility that an observed spectrum can be reproduced by multiple different combinations of model parameters.

Currently, several surface property products like LAI, $C_{ab}$, and fAPAR are derived from remote sensing observations and provided to the user community by space agencies regularly (Baret and Guyot 1991). However, without exception, the retrieval of these properties suffers from ill-posedness, which impacts the quality of the products. Besides the uncertainties in the magnitudes of the retrieved surface properties, the ill-posedness has also caused unrealistic short-term fluctuations, e.g., in the MODIS LAI and fAPAR products (Garrigues et al. 2008).

A priori knowledge about the surface and atmospheric properties can reduce the ill-posedness of a model inversion problem and the uncertainty in the selected solution (Lauvernet et al. 2008; Verrelst et al. 2015). Bare soil reflectance from near surroundings is commonly used as a prior for the background underneath a vegetation canopy (Darvishzadeh et al. 2008). Leaf angle distributions, which strongly affect observed spectra, can be roughly determined according to vegetation types (Zarco-Tejada et al. 2004). In addition to the prescribed a priori knowledge, a priori information about vegetation properties can be obtained from process models like vegetation growth and prognostic phenology models (Fang et al. 2008; Knorr et al. 2010; Savoy and Mackay 2015; Weiss et al. 2000). In principle, process models can provide information about which solutions (e.g., seasonal variation of LAI) are more likely than others. However, these types of models are mainly available for crops. Moreover, many important controlling input and state variables are required in the process models to predict vegetation growth, such as wind speed, air temperature or rainfall. Without accurate values of these variables, the predictions of the vegetation state variables may not serve as proper prior information of model inversion problems. As a result, it turns out to be difficult to establish a general and operational link between process models and RTMs for retrieving vegetation properties from satellite observations, although a few successful cases have been reported (Quaife et al. 2008). In contrast to mechanistic models, empirical vegetation dynamic models with climatology information require less input and sometimes rely on existing measurements or remote sensing products, such as LAI climatology from multi-temporal years of MODIS LAI data (Liu et al. 2014).

The temporal autocorrelation and dependence of land surface and atmospheric properties are promising sources of prior information that have not been well exploited yet. Tobler’s first law of geography states that “everything is related to everything else, but near things are more related than distant things” (Tobler 1970). This law describes the spatial contiguity of objects and is the foundation of the fundamental concepts of spatial autocorrelation and dependence, which have been used in numerous applications in geoscience, notably, for spatial interpolation of land surface and atmospheric properties (Miller 2004; Sui 2004). Tobler’s first law of geography is generally valid without underlying assumptions unless human interferences are involved. The concept of spatial contiguity can be extended to the temporal domain as the temporal contiguity of the state variables of an object or a system. An analogous law is, therefore, established as “every state of an object is related to its states in every other moment, but temporally near states are more related than distant ones”. This law describes an implicit but consonant source of prior information for retrieving surface biophysical variables from time series of satellite observations, namely surface state variables at a certain moment are related to the state variables at previous moments. Similar to the spatial contiguity, the correlation between two states of variables is negatively related to the temporal gap between the two states.

Most algorithms for retrieving surface biophysical variables from the satellite observations exploit the spectral and occasionally the directional variation of the radiometric information and the temporal contiguity of surface and atmospheric properties as constraints in model inversion problems (Lauvernet et al. 2008; Liu et al. 2014; Verrelst et al. 2015; Weiss et al. 2000). Several studies have attempted to incorporate the temporal autocorrelation of land surface and atmospheric properties to constrain model inversion problems and to improve the reliability and consistency of vegetation variables (Samain et al. 2008; Shi et al. 2017; Xiao et al. 2011). The use of vegetation growth models or empirical vegetation dynamic models (Koetz et al. 2005; Liu et al. 2014) is a possible way to provide temporal continuity of some vegetation properties (e.g., LAI). Besides the temporal continuity obtained from vegetation growth and dynamic models, some studies assumed that the vegetation state was not to vary over a temporal window of a few days and found significant performance improvements in retrieving some canopy characteristics (e.g., LAI and leaf angles) by incorporating multi-temporal satellite observations (e.g., Lauvernet et al. 2008). The assumption of invariant vegetation properties used in Lauvernet et al. (2008) only applies to some crop types and small temporal windows. It is not valid in many other cases, such as vegetation at development or senescent stages. The use of vegetation growth models, dynamic models and the assumption of invariant vegetation properties in model inversion problems are examples of applications of temporal continuity, but the natural temporal dynamics of vegetation can vary from the predictions of these models due to some embedding assumptions or violate the (invariant) assumption. Compared with this prior information, a safer temporal constraint would be based on Tobler’s first law of geography in the temporal domain, which is generally valid except for human interferences. Mousivand et al. (2015) recognized the limitation of these assumptions and proposed a prototype of a retrieval algorithm. The proposed algorithm allows incorporating the retrieved vegetation variables from previous observations as prior information while considering the temporal variation of vegetation properties. This is a plausible and promising prototype to utilize the temporal autocorrelation of vegetation and land surface properties to constrain model inversion problems.

However, the focus of Mousivand et al., (2015) study was on the synergy of multi-sensor observations for the retrieval by using a coupled surface-atmosphere model, and the gain of using the temporal continuity as prior information was not fully exploited. Only a simple experiment using measurements on five consecutive days was conducted to illustrate the feasibility of the prototype. Additionally, they utilized the coupled soil-leaf-canopy (SLC) model with the MODTRAN atmospheric RTM to perform the retrieval (Berk et al. 2005).

The coupling of canopy RTMs with MODTRAN appears to be a common practice (Grau and Gastellu-Etchegorry 2013; Laurent et al. 2014; Mousivand et al. 2015; Verhoef et al. 2018; Verhoef and Bach 2003), but MODTRAN is of high computational complexity and difficult to invert. As a result, measurements of atmospheric properties are provided to MODTRAN to simulate the optical coefficients (atmospheric transmittance factors), which are then used to translate top-of-canopy (TOC) to top-of-atmosphere (TOA) observations. In the subsequent retrieval, only the surface properties are obtained by fitting simulated TOA observations with measured ones. This procedure poses limitations since it requires accurate measurements of atmospheric properties. This limitation can be overcome by using simpler atmosphere RTMs. The Soil-Plant-Airshphere Radiative Transfer model (SPART) couples computationally efficient RTMs for soil, vegetation canopies and atmosphere (Yang et al. 2020a). The more efficient atmosphere RTM
makes the whole SPART model easy to invert, allowing estimation of vegetation and atmospheric properties directly from TOA observations by model inversion.

Success in using these temporal constraints requires satellite observations that are frequent enough to ensure the temporal autocorrelation and independence of the attributes between two successive measurements. Moreover, a retrieval framework that makes use of autocorrelation and dependence of land surface and atmospheric properties is required. The number of earth observation satellites in the optical domain has increased spectacularly in the last decades. These satellites provide numerous time series of multi-spectral observations of land surface with high temporal resolutions, such as data from the Moderate Resolution Imaging Spectroradiometer (MODIS) and the Advanced Very-High-Resolution Radiometer (AVHRR) (Combal et al. 2003; Saunders and Kriebel 1988). Additionally, the Ocean and Land Colour Instrument (OLCI) aboard the Sentinel-3 (S3) satellite measures visible and infrared radiance since 19 April 2016, with a revisit time of 1.1 days (Donlon et al. 2012). The high revisit time and the early overpass time (before 11:20 a.m.) enable the monitoring of vegetation over the growing season and limit the problem of clouds. It is evident that, nowadays, we already have access to a large amount of satellite data with a high temporal resolution.

In this study, we aim to establish a retrieval framework that makes use of the temporal autocorrelation and dependence of land surface and atmospheric properties and derive high-level products of OLCI data with a focus on LAI by using the SPART model. The choice of using S3 data is also motivated by the fact that the ESA’s forthcoming FLuorescence EXplorer (FLEX) mission, which is dedicated to global monitoring of photosynthesis, will fly in tandem with S3 (Kraft et al. 2012). For decoupling information of photosynthesis and non-photosynthetic factors from fluorescence observations, leaf biophysical and canopy structural properties need to be determined concomitantly (De Grave et al. 2020), and FLEX mission will employ S3 data to characterize land surface biophysical properties.

2. Materials and methods

2.1. The SPART model

To retrieve the properties of a soil-plant-atmosphere ensemble from TOA radiance observations, one requires a model that links these properties with the TOA observations. In this study, we use the Soil-Plant-Atmosphere Radiative Transfer (SPART) model, which couples three widely used, computationally efficient RTMs and simulates TOA reflectance and radiance in any given direction. In what follows, we provide a brief introduction of the SPART model, and for the details the reader is referred to the original publication (Yang et al. 2020a).

TOA reflectance and radiance observed by a sensor are the outcomes of solar light interaction with the atmosphere, vegetation canopy and soil background. To simulate TOA reflectance/radiance, the SPART model first computes the optical properties of the soil background (i.e., soil reflectance), vegetation and atmosphere layer (i.e., reflectance and transmittance) with three independent sub-models. They are the BSM (Brightness-Shape-Moisture) soil RTM (Verhoef et al. 2018), the PROSAIL canopy RTM (Jacquemoud et al. 2009) and the SMAC atmosphere RTM (Rahman and Dedieu 1994) with some modifications. The three RTMs are further coupled by using the adding method, which calculates the optical properties of an ensemble such as a two-layer system or a surface-layer system (Verhoef 1985; Yang et al. 2020b).

BSM simulates the isotropic soil reflectance and requires the inputs of soil brightness ($B$), soil moisture ($SM_{s}$), and two spectral-shape related parameters ($\phi$ and $\lambda$) (Table 1). This model is based on an empirical reflectance model of dry soil (Jiang and Fang 2019; Verhoef et al. 2018) and incorporates the effects of soil moisture on soil reflectance by using the water-film approach (Yang et al. 2020a).

PROSAIL couples the PROSPECT leaf RTM and SAIL (Verhoef 1985, 1984) canopy RTM and simulates anisotropic canopy reflectance factors. PROSPECT and SAIL are the most widely-used leaf and canopy RTMs, respectively, and have continuously been revised and improved. In the SPART model, PROSPECT-D is used (Féret et al. 2017), which is a new version of PROSPECT and which is re-calibrated to include the absorption of chlorophylls, carotenoids and anthocyanins pigments. It requires the inputs of the content of these leaf pigments, senescence materials, water content, dry matter and leaf internal structure (Table 1). A recent version of PROSPECT called PROSECT-PRO allows for proper decomposition of leaf dry matter into nitrogen-based proteins and carbon-based constituents, but this decomposition mainly affects leaf reflectance from 2000 to 2500 nm, and has little effects on OLCI reflectance (i.e., from 400 to 1200 nm). The version of SAIL considering the hotspot effects, namely SAILH (Verhoef 1998), is implemented in SPART. This model requires the inputs of the leaf inclination distribution and LAI and a hotspot parameter (i.e., the ratio of leaf size to vegetation height). It up-scales leaf optical properties (i.e., leaf reflectance and transmittance) to canopy optical signals by considering the canopy architecture. PROSAIL is a 1D model assuming a vegetation canopies as a turbid medium and is more suitable for crop and grass than forest canopies. Nevertheless, 1D and 3D models tend to converge in simulating canopy reflectance at coarser spatial resolutions (typically beyond 100 m) as the role of the detailed 3D canopy structure on reflectance becomes less crucial (Widlowski et al. 2005). The PROSAIL model appears suitable for satellite applications due to its simplicity and reasonable accuracy.

SMAC is based on the 6S atmosphere RTM (Vermote et al. 1997) but reduces model complexity and computation time by simplifying the computation of the process variables, mostly by using semi-empirical fitting functions (Proud et al. 2010). It requires the same seven inputs

Table 1: Main inputs of the SPART model, and their physically plausible ranges, initial values and relaxation time used in the retrieval algorithms.

| Parameter | Description | Unit | Range | Initial value | Relaxation time (days) |
|-----------|-------------|------|-------|---------------|-----------------------|
| $B$       | Soil brightness | –    | [0.0,9] | 0.5           | 2                     |
| $\phi$    | Soil spectral latitude | Degree | [−30,30] | 0             | 60                    |
| $\lambda$ | Soil spectral longitude | Degree | [80,120]  | 100           | 60                    |
| $SM_{s}$  | Soil moisture volume percentage | – | [5,55]  | 20            | 2                     |
| $C_{ab}$  | Chlorophyll a and b content | $\mu g$ | $cm^{-2}$ | [0.80]       | 40                    |
| $C_{on}$  | Dry mass per unit leaf area | – | g cm$^{-2}$ | [0.0,02]     | 0.01                  |
| $C_{w}$   | Equivalent water thickness | cm  | [0,0.1] | 0.02          | 30                    |
| $C_{s}$   | Senescent materials | – | [0,1]  | 0             | 30                    |
| $C_{c}$   | Carotenoid content | $\mu g$ | $cm^{-2}$ | [0,30]       | 10                    |
| $N$       | Leaf internal structure parameter | – | [1,4]  | 1.5           | 60                    |
| LAI       | Leaf area index | $m^2$ | $m^2$ | 3             | 30                    |
| LDFa      | Leaf inclination determination parameter a | – | [−1,1] | 0.35          | 30                    |
| LDFb      | Leaf inclination determination parameter b | – | [−1,1] | 0.15          | 30                    |
| AOT550    | Aerosol optical thickness at 550 nm | – | [0,2]  | 0.325         | 2                     |
| $U_{O3}$  | Ozone content | cm-atm | [0,0.8] | 0.35         | 60                    |
| $U_{vap}$ | Water vapour | g cm$^{-2}$ | [0,8,5] | 1.41        | 2                     |
of 6S, which are the sun-observer geometry parameters (i.e., $\phi_s$, $\theta_s$, and $\phi_0$), aerosol optical thickness at 550 nm ($AOT_{550}$), ozone content ($U_{O3}$), water vapour content ($U_{H2O}$) and air pressure ($P_a$).

### 2.2. Synthetic dataset

We first generated a synthetic dataset consisting of 35 different scenarios by using the SPART model. The physical properties of soil-vegetation-atmosphere systems were user-defined and known, and the corresponding TOA radiance of the scene in a given viewing direction was simulated with SPART. The synthetic dataset comprised soil properties, leaf properties, canopy structure, sun-observer geometry and the corresponding TOA radiance. Afterwards, some artificial measurement uncertainties were added to the simulated TOA radiance to obtain realistic noise-contaminated OLCI observations. In this study, we took a corn canopy during a representative growing season to generate virtual scenarios. The complete dataset consisted of the model input and output, which were respectively defined and generated in the following steps.

First, the corn canopy was characterized by using the PROSAIL model parametrization (presented in Table 1). Taking representativeness and simplicity of the scenarios into account, we considered the seasonal variation of two key vegetation parameters, leaf area index and chlorophyll content in their generation. The seasonal LAI and $C_{ab}$ assigned in the virtual scenarios are presented in Fig. 1. A generic seasonal dynamic of LAI was extracted from WOFOST (Van Diepen et al. 1989), which is a widely-used crop growth model. The (northern hemisphere) growing season started on the day of the year (DOY) 170 and ended on DOY 274. The seasonal dynamic of leaf chlorophyll content is not directly predicted by WOFOST, but was generated based on the in situ measurements from the existing literature, e.g., Yan et al. (2017) and Ciganda et al. (2008). The other vegetation parameters were kept constant throughout the growing season. Generic corn leaf data provided in the PROSPECT-3 documentation (Elseikih and Sommer 1980) was used, of which the equivalent water thickness $C_w = 0.009$ cm, dry matter $C_{dm} = 0.0021$ g cm$^{-2}$, and mesophyll structural parameter $N = 1.5$. These leaf properties may vary across a growing season, but their seasonal variation is usually less pronounced than that of LAI and $C_{ab}$. Leaf water thickness, however, can change rapidly in response to variation in soil moisture, but its impact on reflectance is mainly in short-wave infrared, which is out of the spectral region of the OLCI data. Additionally, a spherical leaf angle distribution (i.e., $LIDFa = -0.35$, $LIDFb = -0.15$) was assumed for the corn canopy in the growing season. The parameters for both soil properties and atmospheric properties were set to the default values in the SPART model (Table 1).

Second, we simulated nadir-viewing TOA radiance spectra of the canopy once every three days at noon by using the SPART model. The sensor for simulation was set to OLCI, which determines the spectral configuration (spectral range and resolution) of the radiance. The temporal resolution of three days was chosen according to the temporal resolution of OLCI observations before S3-B became available. The sun position (i.e., solar zenith and azimuth angles) was computed according to the date, time and geographic coordinates, which was set to the Economic and Environmental Enhancement (OPE3) site in Beltsville, MD, USA (39.0306°N 76.8454°W, UTC-5).

Third, in order to mimic satellite measurements of TOA radiance, two types of errors were added to the simulated reflectance. The first type was a linear function of the signal (the simulated radiance) mimicking the measurement noise. The effective noise model proposed by Verhoef et al. (2018) was used, where the standard deviation of the measured noise radiance is given as:

$$\Delta L = \sqrt{aL + b}$$

where $L$ is the measured/simulated TOA radiance in units of mW m$^{-2}$ nm$^{-1}$ sr$^{-1}$, $a$ and $b$ are constants that are specific for the given sensor, which were set as $1.136 \times 10^{-4}$ mW m$^{-2}$ nm$^{-1}$ sr$^{-1}$ and $2.724 \times 10^{-4}$ mW m$^{-2}$ nm$^{-2}$ sr$^{-2}$, respectively, according to Verhoef et al. (2018). The second type of errors was added to account for larger measurement uncertainties caused by unfavourable observational conditions, such as dense clouds and unstable weather. Clouds normally enhance the visible reflectance. Therefore, we randomly selected three days in the growing season: DOY179, 209 and 225, and added 10% artificial errors to the simulated radiance at the visible bands. Although two types of errors were added to reproduce remote sensing signals observed by satellites, the discrepancy between model simulated and satellite measured signals is expected due to, for example, model representation and parameterization errors (Berger et al. 2018; Widlowski et al. 2005). Therefore, it is necessary to test the algorithms with real satellite data besides the synthetic dataset.

### 2.3. OLCI satellite dataset

Besides the synthetic dataset, a dataset acquired by the OLCI sensor was used. OLCI measures reflected radiance of earth surface covering the spectral range from 403 to 1040 nm with bandwidths ranging from 3.75 to 40 nm with 21 bands. The band locations of OLCI are provided in Fig. 1b. Its swath width is 1270 km with a spatial resolution around 300 m at nadir. The temporal resolution of OLCI is approximately 2–4 days and has doubled since the launch of S3-B and the revisit time is less than two days at the equator after 25 April 2018. Level-1 OLCI TOA radiance data are freely available. They were downloaded from Copernicus open access hub (Sentinel 2018) from April 26, 2016 (the earliest possible) to September 1, 2019.

TOA radiance observations from OLCI (level-1 products) of several

---

**Fig. 1.** Representative seasonal dynamics of leaf area index (LAI) and leaf chlorophyll content (Cab) of corn canopies used in the synthetic dataset, and the corresponding TOA radiance on three days (DOY170, DOY210 and DOY260). The grey line segments perpendicular to the abscissa axis in b) indicate the central wavelengths of the 21 bands of OLCI.
The Harvard Forest, which is an ecological research area of 3000 acres in Massachusetts. The site is predominantly red oak and red maple, with trees and a herbaceous layer. Three annual herbaceous plants, namely grasses, forbs and legumes, which are green and active from October to the end of May, are predominant at the site. The US-Ha1 site is located in Bondville in the central US and is covered with temporary crops followed by harvest and a bare soil period. The main crops are corn (Zea mays L.) and soybean (Glycine max L.) (Chen et al. 2018).

The ES-LMa site is a Mediterranean tree-grass ecosystem located in southwestern Spain. This is a managed savannah consisting of sparse trees and a herbaceous layer. Trees (mainly Quercus ilex L.) are separated by distances of \( \sim 18.8 \pm 5 \) m resulting in the fractional cover \( \sim 20\% \) (Pacheco-Labrador et al. 2016). Three annual herbaceous plants, namely L. cornutum, excelsior, and P. Yang et al.

Once we know the effects of a change of the parameter vector on TOA radiance and the difference of TOA radiance from model simulation and measurements, the algorithm updates the parameters to reduce the difference of TOA radiance, which was the cost function. The model was iteratively executed, and iteration stopped when the improvement of the cost function \( f_o \) was less than \( 10^{-3} \), which takes around 35 iterations when the standard initial guesses (in Table 1) are applied according to our experiment. Note that the number of iterations reduces to less than 10 if the temporal continuity is considered because more reasonable initial guesses can be obtained from the retrievals of the previous measurements.

### 2.4. Retrieval algorithm for independent observations

We used the numerical optimization method to retrieve the parameters from both synthetic and satellite datasets. The numerical optimization method aims at minimizing a cost function, which quantifies the differences between measured and simulated signals by successive changes of the input parameters, considering possible prior information.

In the first algorithm, the retrieval of vegetation and atmospheric properties from TOA radiance was conducted on each day independently using no prior information at all. The cost function (Eq. (2)) only depends on the difference between measured and modelled radiometric data and is expressed as:

\[
f_o(i) = (L_{\text{mod}}(i) - L_{\text{meas}}(i))^T \sigma_i^{-1} (L_{\text{mod}}(i) - L_{\text{meas}}(i))
\]

where \( L_{\text{meas}} \) and \( L_{\text{mod}} \) are the observed and modelled TOA radiance. Note that data from the synthetic dataset are considered as observations in the algorithm, although they are simulated with a model. The superscript 'T' denotes the transpose of a matrix, and the symbol 'i' denotes the index of the measurements in the time series. \( \sigma_i \) is the standard deviation of the observations describing the uncertainties in the prior observations, which were estimated by using the effective noise model according to Eq. (1). Hereafter, STDalgo is used to denote this standard retrieval algorithm.

Vegetation properties were determined by the radiance spectrum at each moment. To perform the numerical optimization, we used the function ‘lsqnonlin’ of the optimization toolbox of Matlab R2017a, selecting a Trust Region algorithm for updating parameter values within the ranges shown in Table 1. The update of model parameters was based on the Jacobian matrix of the model, which quantifies the magnitude and direction (i.e., increase or decrease) of a change in model input parameters on the model output. In this case, it quantified the change in TOA radiance induced by soil, vegetation, and atmospheric properties.

### 2.5. Retrieval algorithm with the use of temporal autocorrelation

In the second algorithm, apart from the radiometric observations, the existing correlation of soil, vegetation and atmospheric properties between two or more successive measurements were used to constrain the retrieval. The key of this algorithm is to employ the estimated vegetation properties from earlier measurements in a time series as a memory of the parameters at the current moment. According to this idea, a general form of the cost function of our algorithm for time-series observations can be expressed as:

\[
f(i) = f_o(i) + f_p(i) \\
= (L_{\text{meas}}(i) - L_{\text{mod}}(i))^T \sigma_i^{-1} (L_{\text{meas}}(i) - L_{\text{mod}}(i)) \\
f_p(i) = \sum_{m=1}^{n} [f(i-m) - f_{\text{mod}}(i)] [P(i) - P(i-m)] \sigma_p^{-1}(i-m) [P(i) - P(i-m)]
\]

which not only includes \( f_o(i) \) that quantifies the difference between measured and modelled radiance, but also \( f_p(i) \) that quantifies the changes between the prediction of the vegetation properties at the current moment and the estimated parameters from the previous measurements. Hereafter, TSalgo is used to denote this time-series retrieval algorithm.

In Eq. (3), \( f_p \) is an additional constraint to the model inversion problem. The vector \( P(i) \) is the prediction of the parameters at the current moment, and \( P(i-m) \) is the estimated parameters of the previous \( m \)th measurement. The previous \( n \) (i.e., \( n \geq 1 \)) measurements are used to constrain the current retrieval. The quantity \( \sigma_p(i-m) \) is the standard deviation of \( P(i) - P(i-m) \) describing the uncertainties in the prior values. The condition \( f(i-m) < f_{\text{mod}}(i) \) is a boolean function that determines whether the retrieved parameters from the \( m \)th observation will be used or not, depending on the reliability or successfulness of previous retrievals. Only retrievals with residuals less than a given threshold (i.e., \( f_{\text{th}} \)) will be used to constrain the current retrieval. This reduces the error propagation from the previous unsuccessful observations (e.g., due to clouds) to the following retrievals. The quantities \( n \) and \( f_{\text{th}} \) can be configured according to land covers and temporal resolutions of satellite observations. In practice, the value of \( f_{\text{th}} \) can be determined based on the residuals of the spectral fitting when TSalgo is applied (i.e., \( f_o \)). It is worth noting that because the consideration of temporal continuity in the cost function results in a slight increase of the difference between modelled and measured spectra, \( f_o \) in TSalgo should be slightly higher than the spectral residuals (i.e., temporal mean \( f_o \) in

### Table 2

Information about the study sites.

| Site ID | Site name       | Country      | Lat/Lon (degree) | Land cover         |
|---------|-----------------|--------------|-----------------|--------------------|
| CH-Lae  | Lageren         | Switzerland  | 47.48/8.37      | MF (Mixed forest)  |
| US-B01  | Bondville       | US           | 40.01/–88.29    | CRO (Cropland)     |
| ES-LMa  | Majadas de Tietar | Spain       | 39.94/–5.77     | SAV (Savannahs)    |
| US-Ha1  | Harvard Forest  | US           | 42.54/–72.19    | DBF (Deciduous broadleaf forest) |
Eq. (2)) in STSalgo. The number of days of measurements used to constrain the current retrieval (i.e., n) is an empirical parameter. In our experiment, n and \( f_{\text{mean}} \) were set to four (days) and 2.5 \( \times 10^4 \), respectively.

It is clear that the uncertainties of the prior information are very important. In TSalgo, the estimated parameters from previous moments are used as priors, but changes in the parameter values between consecutive time steps should be permitted and accounted for by using \( \sigma_p \). In this respect, the uncertainties \( \sigma_p \) can be considered to represent the resistances to change for the parameters. They depend on 1) the time gap (\( \Delta t \)) between observations of the same target and 2) the relaxation time \( (t_p) \) of the parameters that quantifies how fast the parameters can vary (Mousivand et al. 2015).

Naturally, most surface properties change gradually between two successive observations and the resistance against change decreases with the temporal gap between observations of the same target. The rates of change of different surface properties are, however, not the same. For instance, the background dry soil spectrum is not very likely to change over the year and can be given a low resistance to change. Most vegetation parameters, like the LAI and leaf chlorophyll content, will change only slowly as a function of time, while soil moisture in the topsoil may vary from day to day and should be given a high resistance to change. Considering the above-mentioned factors, the resistance to change (Eq. (4)) was estimated as a function of the physically possible or biologically plausible range of the respective parameter (\( P_{\text{max}} - P_{\text{min}} \)) (Table 1) and of the temporal gap between two observations as follows:

\[
\sigma_p (i-m) = \frac{P_{\text{max}} - P_{\text{min}}}{\sqrt{1 + e^{-\Delta t / t_p}}} \tag{4}
\]

where \( \Delta t = t(i) - t(i-1) \) is the time gap between the current observation and the previous \( m \)th observation and \( t_p \) is a relaxation time, which is parameter-dependent. Specifically, \( t_p \) is defined as the time it would take for the standard deviation of the change of a parameter between two moments (i.e., \( P(i) - P(i-1) \)) to become equal to the one corresponding to a uniform distribution covering the full possible range of that parameter. The standard deviation of parameters with a uniform distribution over the given interval can be computed as \( \frac{1}{\sqrt{12}} \) (\( \approx 0.3 \)) of \( P_{\text{max}} - P_{\text{min}} \), which is about 30% of the full possible range (Emad Noujeim 2019).

The time-series retrieval algorithm (TSalgo) includes a strong resistance to change for certain parameters by assigning a long relaxation time if it is known that these parameters must be almost constant in time. In practice, the parameters in the SPART model can be classified into three categories based on their resistance to change or the relaxation time. The first category includes soil moisture \( (\text{SM}_{\text{e}}) \) (Richter et al. 2012), soil brightness \( B \) and aerosol optical thickness \( (\text{AOT}_{\text{TOC}}) \) and water vapour \( (U_{\text{TOC}}) \). These parameters should be given a low resistance to change (e.g., in our study, the relaxation time was set to two days, \( t_p = 2 \) days). The second category includes LAI, leaf water content \( (C_{\text{wa}}) \), LIDF, leaf chlorophyll content \( (C_{\text{ch}}) \) and carotenoid content \( (C_{\text{ca}}) \). A medium resistance to change should be given to these parameters (e.g., \( t_p = 30 \) days). The third category includes soil spectral-shape parameters \( (\alpha \text{ and } \lambda) \) and dry matter \( (C_{\text{dm}}) \) and leaf structural parameter \( (N) \) and Ozone content \( (U_{\text{O3}}) \). A long relaxation time should be given (e.g., \( t_p = 60 \) days). It should be emphasized that the relaxation times of most parameters are based on expert knowledge, but some of them should be changed according to the vegetation type (Mousivand et al. 2015). If the relaxation time of the parameter vector is too long (i.e., high resistance), one will obtain large residuals in radiometric data fitting \( (f_g(i)) \) for most of the days. This is because the parameters are forced to be close to the prior coming from the previous days, and the measured reflectance on the current day contributes little to the cost function.

The relaxation times were assigned to each variable to characterize changes of land surface and atmospheric properties due to natural processes. For land surface changes caused by human interferences, such as cutting of grassland and deforestation, the retrieved variables from previous measurements are inappropriate prior information for the current moment. This can result in more iterations in the optimization and a larger value of the cost function (Eq. (3)). On the one hand, temporal autocorrelation does not improve the accuracy of the retrieved variables when human interferences exist. On the other hand, by looking at the number of iterations and the residuals after the optimization, it is possible to detect the changes caused by human interferences. An alarm can be included when the number of iterations exceeds a threshold. However, the detection of human interferences requires a different experiment set-up, and we focused on changes due to natural processes in this study.

### 2.6. Validation and evaluation

In the numerical experiment with the synthetic dataset, we validated the retrieval algorithms by comparing the estimated LAI and leaf chlorophyll content with the values assigned during the dataset generation. For the OLCI satellite, the estimated LAI from the OLCI observations was benchmarked with the MODIS LAI products MCD15A3H version 6 (Myneni et al. 2002) for the four sites with the focus of seasonal variation. Furthermore, for the forest sites, the estimated LAI was compared with some published field measurements of LAI.

The MCD15A3H LAI product is a 4-day composite data with 500-m pixel size, which is freely available on the NASA data repository and which was downloaded via the Application for Extracting and Exploring Analysis Ready Samples (AppEEARS, https://lpdaacsvc.cr.usgs.gov/appèears). The operational algorithm of this MODIS product is based on atmospherically corrected TOC reflectance. A look-up-table (LUT) method is used to achieve inversion of the radiative transfer problem, and a back-up method based on empirical relations between the normalized difference vegetation index (NDVI) and LAI, together with a biome classification map, is used when the LUT method fails to localize a solution (Knyazikhin 1999; Myneni et al. 2002). MODIS LAI products can have low accuracy because of (1) uncertainties in the input data, such as errors in remote sensing reflectance due to atmospheric effects and clouds; (2) model uncertainties and problems associated with ill-posed retrieval; and (3) errors in the ancillary information (e.g., land cover type and atmospheric properties) (Xu et al. 2016). Therefore, MODIS LAI products were used for plausibility checks rather than a quantitative validation.

The LAI measurements at the CH-Lae site were collected from Paul-Limoges et al. (2017). The overstory leaf area index (LAI) was measured using an LAI-2000 (LICOR Inc., Lincoln, NE, USA) once a month during the 2015 growing season. Understory LAI was estimated from the understory biomass measurements, which were taken once a month during a part of the 2014 growing season and during the full 2015 growing season. A simple empirical approach proposed by Ernst (1979) was applied to estimate understory LAI from the biomass measurements. The LAI measurements at the US-Ha1 site were collected from Yang et al. (2017). The measurements were taken at a daily time step during spring and autumn and biweekly intervals during the 2014 growing season (~May–October) by using an LAI-2000 Plant Canopy Analyzer (LI-COR, Inc., Lincoln, NE, USA). Although the acquisition time of the field measurements is in 2015 while the OLCI data is available since 2016, the seasonal cycle of canopy LAI of mature natural forests is generally similar among years, and it is reasonable to evaluate the plausibility of the retrieved results by using the field measurements. The LAI of crop-lands, however, varies with the crop plantation, e.g., corn and wheat obviously have different phenology.

We also compared the retrieved atmospheric aerosol optical thickness \( \text{AOT}_{\text{TOC}} \) with the products from ECMWF (European Centre for Medium-Range Weather Forecasts). The total aerosol optical depth at 550 nm \( (\text{AOT}_{\text{TOC}}) \) values were downloaded from near-real-time Coper- nicus Atmosphere Monitoring Service (CAMS, https://apps.ecmwf.int/datasets) dataset at surface level. The data has 3-hourly resolution, thus the value at the time of OLCI overpass was acquired by temporal...
3. Results

3.1. Retrievals from the synthetic dataset

The retrieved LAI and leaf chlorophyll content $C_{ab}$ in the growing season, together with the values assigned in the synthetic scenes, are presented in Fig. 2. The general patterns in the seasonal development of LAI and $C_{ab}$ are reasonably well captured by the retrieved values for both retrieval algorithms. Specifically, the estimated values increase from 0.5 to 3.8 m$^2$ m$^{-2}$ for LAI and 30 to 60 µg cm$^{-2}$ for $C_{ab}$, maintain at high values for about 70 days and decrease rapidly at the senescent stage.

The simple numerical experiment shows evidently that the use of the temporal continuity of the surface and atmospheric properties in the retrieval scheme improves the accuracy of the estimation of LAI and $C_{ab}$. The prior information generated from the previous observations reduces the ill-posedness of the model inversion problem by comparing the results for TSalgo with those for STDalgo. The errors in the LAI and $C_{ab}$ estimation are as large as 1.3 m$^2$ m$^{-2}$ and 17.3 µg cm$^{-2}$, respectively, when every spectral observation is inverted independently (i.e., the standard retrieval algorithm). In contrast, the errors reduce substantially and are less than 0.6 m$^2$ m$^{-2}$ for LAI and 3 µg cm$^{-2}$ for $C_{ab}$ when prior information is used.

On the days when the artificial errors are added, the $C_{ab}$ estimates differ considerably from the values assigned to the synthetic scenarios (e.g., on DOY 225), but the difference between the assigned and estimated $C_{ab}$ are largely reduced when TSalgo is used. The additional errors in visible bands mimicking the effects of clouds appear to cause an underestimation of LAI with the time-series retrieval scheme on DOY209, and an overestimation on DOY225. However, the error of the LAI retrievals is not significantly greater on the days when we add the artificial noise, but the root-mean-square error (RMSE) between modelled and measured TOA reflectance after optimization is substantially higher on these days as shown in Fig. A1. It is worth noting that residuals of spectral fitting are greater on most of the days when TSalgo is applied, because in STDalgo the cost function (Eq. (2)) just includes radiometric information while in TSalgo a compromise between reproducing the measured data and considering temporal continuity is made.

Fig. 3 displays considerable deviations between the estimated and assigned values for the soil parameter $\phi$, leaf structure parameter $LIDFa$, canopy structure parameter $LIDFb$ and atmospheric $AOT_{550}$ when the observations at different time are treated independently. Because the model inversion problem is generally ill-posed, the under/over-estimation of LAI and $C_{ab}$ is very likely associated with the poor estimation of the other parameters. The use of the temporal contiguity of the land surface and atmospheric properties improves the accuracy of the estimation. At the first several measurements, there are still considerable errors in the estimated $LIDFb$ and $AOT_{550}$, because the available prior information from the previous is null for the retrieval from the first observations. However, the errors rapidly reduce when more observations are involved in constraining the model inversion problem. As a result, the estimates of the model parameters converge to their assigned values. The soil parameter $\phi$ is one exception. The errors between the estimated values and the assigned values are small at the early season when the LAI is low, but they increase with the development of the corn canopy. TOA spectra are sensitive to soil reflectance when the vegetation coverage (expressed as LAI) is low, and the impact of the soil reflectance decreases with the increasing vegetation coverage.

3.2. LAI from the OLCI satellite dataset

Next, we examine the performance of the retrieval schemes with and without using the temporal continuity of land surface and atmospheric properties for the satellite observations (i.e., STDalgo vs. TSalgo). The retrievals of LAI of the four sites, together with the MODIS LAI products are shown in Fig. 4, followed by the comparison of the retrieved LAI from the satellite observations with the field measurements for two forest sites (Figs. 5 and 6).

Fig. 4 demonstrates that solely utilizing single radiometric observations for retrieving LAI can result in a large number of unrealistic fluctuations (i.e., the spikes in the blue lines in Fig. 4). Although this kind of fluctuations also exists when TSalgo is applied, the incorporation of the estimates from the previous observations suppresses the LAI changes between successive acquisitions. As a result, the LAI retrievals with the time-series retrieval approach are smoother. The fluctuations also occur for the MODIS LAI, and the frequency of the occurrences appears to fall in between the results from STDalgo and TSalgo.

At all the four sites (i.e., CH-Lae, US-Bo1, ES-LMa and US-Ha1), the estimated LAI with both STDalgo and TSalgo, as well as the MODIS LAI, present similar seasonal patterns, which vary among the sites due to the different plant phenology. The mixed forest (CH-Lae) site has pronounced seasonality with low LAI in the winter (from December to March) and high LAI in the summer (from May to September). The estimated LAI values from OLCI and the MODIS LAI are close to zero from every December to March of the next year, but increase dramatically after March and peak in July (Fig. 4a). The decrease of LAI from July to December is almost symmetric to the increase from March to July. Several sharp spikes are observed in the MODIS LAI product, such as on August 9, 2017, September 6, 2017, June 18, 2018, and July 28, 2019. The spikes in the MODIS LAI mainly occur between June and September, while in the retrieved LAI with the standard algorithm the spikes arise throughout the season. These sharp changes are physically impossible and must be caused by error in the MODIS LAI product and by STDalgo. In comparison with the MODIS LAI and the retrievals with STDalgo, the spikes in the retrievals with TSalgo are much less frequent and more subtle. RMSE between modelled and measured TOA radiance after optimization is usually greater in winter than in summer (Fig. A2), most likely due to greater model representation errors for the land.

![Fig. 2. Estimates of LAI and chlorophyll content ($C_{ab}$) of a synthetic corn canopy for a growing season by using the standard retrieval algorithm (STDalgo) and the time series retrieval algorithm (TSalgo), and the values assigned to the two variables in the synthetic scenes. The days with additional artificial errors in reflectance are indicated with arrows.](image-url)
surface in winter when vegetation coverage is lower, and soil and/or snow are observed.

In comparison with the CH-Lae site, the seasonal pattern of LAI in the cropland (US-Bo1) site shows a more profound growing season from April to October, although at both sites the LAI estimates from OLCI and the MODIS LAI are greatest in July (Fig. 4b). Furthermore, the increase in LAI from April to July and its decrease from July to October is more dramatic. The MODIS LAI and LAI estimated with TSalgo appear to be close to each other, except that the temporal resolution of the estimated LAI from OLCI is substantially higher than the MODIS LAI. Unrealistic dips can be found in both the MODIS LAI and the retrievals from OLCI at the end of July in 2017. Nevertheless, both of them are significantly smoother than the estimated LAI with STDalgo.

At the savannah (ES-LMa) site, in addition to the lower values of the LAI estimates and the MODIS LAI, the seasonal pattern is also different from the CH-Lae and US-Bo1 sites (Fig. 4c). The estimated LAI from OLCI TOA radiance and the MODIS LAI peaks in May, which is about two months earlier than the mixed forest and cropland sites. The MODIS LAI and the retrievals with the time-series approach are close to each other in terms of magnitude and seasonal variation. In contrast, the retrievals from STDalgo are clearly problematic, especially between July and December. A number of large over-night changes occur in the estimated LAI with the standard algorithm.

At the deciduous forest (US-Ha1) site, the seasonal patterns of the MODIS LAI and the LAI estimates from OLCI TOA radiance are similar to those at the CH-Lae site, but the magnitude of LAI is slightly lower at the US-Ha1 site than the CH-Lae site by about one m²m⁻² (Fig. 4d).

The field measurements of LAI were collected two years (i.e., in 2015) earlier than the first acquisition of the OLCI data shown in this study and are not ideal for direct validation of the retrieved LAI from the OLCI data. However, as shown in Fig. 4, the LAI of the mature forest is relatively stable over the years, and the field measurements can be used for plausibility checks. Hence, we compared the average of the estimated LAI of the CH-Lae and US-Ha1 forest sites with the field measurements of LAI collected in 2015. Fig. 5 reveals that both the magnitude and seasonal variation of LAI match well between the estimates and the measurements. At the CH-Lae site, we notice that the field measurements of LAI are higher than the estimates from the OLCI data on DOY 289 and DOY 345. The large heterogeneity of the mixed site may have caused the difference. The field measurements obviously cover a smaller footprint than the satellite observations.

Fig. 6 shows that the improvements of the retrieval accuracy by incorporating the temporal continuity of land surface and atmospheric properties in the model inversion are significant. When no constraint is applied to the retrieval from the OLCI observations (Fig. 6b), the R² values between measured and estimated LAI are only 0.46 and 0.28 for the CH-Lae and US-Ha1 sites, respectively. The RMSEs of the estimated LAI are as high as 1.58 m²m⁻² for CH-Lae and 1.10 m²m⁻² for US-Ha1. The accuracy of the MODIS LAI is not satisfactory either (Fig. 6a). The RMSEs of the estimated LAI are 1.29 m²m⁻² for CH-Lae and 0.82 m²m⁻² for US-Ha1 compared with the ground measurements. In contrast, a substantial improvement in LAI estimation for the US-Ha1 site is achieved by using the prior information (Fig. 6c). The RMSE for US-Ha1 is 0.44 m²m⁻², which is 47% and 60% lower than the RMSEs of the MODIS LAI and the retrieved LAI by using STDalgo. The improvements are also apparent if the two sites are considered as a whole. The overall RMSE is 0.80 m²m⁻², which is about 13% improvement compared with the MODIS LAI (RMSE = 0.92 m²m⁻²) and 40% compared with the LAI estimates with only radiometric data (RMSE = 1.33 m²m⁻²). The accuracy of LAI estimation is considerably lower in the CH-Lae site than the US-Ha1 site, most likely due to the stronger heterogeneity of the mixed forest than the deciduous forest.

### 3.3. Other vegetation properties from the OLCI satellite dataset

Next, we examine the other land surface and atmospheric properties retrieved from OLCI TOA observations at the four study sites. Because the retrieval of vegetation parameters is only meaningful and reliable if

---

**Fig. 3.** Estimates of soil parameter \( \varphi \), leaf structure parameter \( N \), canopy structure parameter \( LIDFa \) and atmospheric optical parameter \( AOT_{550} \) of a synthetic corn canopy for a growing season by using the standard retrieval algorithm (STDalgo) and the time-series retrieval algorithm (TSalgo), and the values assigned to the four variables in the synthetic scenes.
vegetation is present during the data acquisition, we selected days with retrieved LAI $> 0.5 \text{ m}^2\text{m}^{-2}$ for the evaluation of the vegetation parameter $LIDFa$ and $C_{ab}$. Fig. 7 shows the seasonal variation of LAI, $LIDFa$ and $C_{ab}$ taken as the average of the values for three growing seasons, i.e., the periods with LAI $> 0.5 \text{ m}^2\text{m}^{-2}$. According to the estimated LAI, vegetation is always present in the savannah site (ES-LMa) and the growing season of the vegetation at the cropland site (US-Bo1) is much shorter than at the two forest sites (CH-Lae and US-Ha1). Furthermore, the peak of LAI at the forest and cropland sites occurs about two months later than the savannah site.

Figs. 8 and 9 depict the retrieved values of $LIDFa$ and $C_{ab}$, respectively, for both the STDalgo and TSalgo for the three years separately. These two figures show that the estimated values of the parameters by using TSalgo are more robust, smoother and stable than those estimated by using STDalgo. In what follows, we focus on the estimated parameters with TSalgo.

The retrieved $LIDFa$ results vary among the sites (Figs. 7c, 8a-d). According to the parameterization of the SAIL model (Verhoef 1998), $LIDFa$ prominently determines the canopy leaf angle distribution and is directly linked with average leaf angle (ALA) as $ALA = 45 - 360 \times LIDFa/\pi^2$ (Yang et al. 2019). Four most common canopy types and their corresponding $LIDFa$ are given as: uniform ($LIDFa$ = 0), planophile ($LIDFa$ = 1), erectophile ($LIDFa$ = −1), and spherical ($LIDFa$ = −0.35). In the forest sites (i.e., CH-Lae and US-Ha1, Fig. 8a and d), the $LIDFa$ values suggest that the canopies are close to spherical canopies, of which $LIDFa = −0.35$. Small seasonal variation is observed at the early and later growing season. The canopies in the ES-LMa site appear to be erectophile, of which $LIDFa = −1$ (Figs. 7c and 8c). The leaves in the canopies at the cropland site (US-Bo1) are more horizontal than those at the other three sites and the leaf angles in this site are uniformly distributed during the growing season, according to the retrieved $LIDFa$ (Figs. 7c and 8b).

Similar leaf chlorophyll content retrievals are found at the two forest sites (Figs. 8b, 9a and d) which complies with the similar $LIDFa$ retrievals at these sites (Figs. 7c, 8a and d). However, a slight temporal shift of twenty days is noticed at the two sites (the blue and gold lines in Fig. 7b) and this also occurs in the retrieved LAI (Fig. 7a). The values of $C_{ab}$ comply with the change of the estimated LAI in the cropland. This is also true at the savannah site except for the time period from DOY 300 to DOY365 when the LAI is low but the estimated $C_{ab}$ increases dramatically. It is worth noting that the retrieval of leaf pigment content (e.g., chlorophyll) is less reliable if the vegetation cover or LAI is low. This may explain the mismatch of LAI and leaf $C_{ab}$ at the savannah site from DOY 300 to DOY365 (the green line in Fig. 7b), as well as the very high estimated $C_{ab}$ values at the two forest sites early in growing seasons before DOY 100 (the blue and gold lines in Fig. 7b).

### 3.4. AOT from the OLCI satellite dataset

Since we used a shorter relaxation time of two days (Table 1), one can expect larger temporal variations in retrieved AOT than in the vegetation parameters. Fig. 10 shows that this is, indeed, the case, but the estimated $AOT_{550}$ is nevertheless more stable when the temporal continuity of the land surface and atmospheric properties is considered in the retrieval. The estimated $AOT_{550}$ from the OLCI observations shows clear seasonal cycles in all the four sites, with higher values in the
summer than in the winter for both retrieval algorithms. In three of the four sites, this seasonal pattern is also visible in the ECMWF AOT data product used as a reference here. Fig. 10 a shows that the AOT retrievals for CH-Lae match well with ECWMF product in terms of both magnitude and seasonal variation. In contrast, at the US-Ha1 site, the ECMWF data are low throughout the season (<0.3), whereas the time series retrieval from OLCI shows a seasonal cycle similar to that of the other sites (Fig. 10 d). In the US-Bo1 (Fig. 10 b) and ES-LMa (Fig. 10 c) sites the ECMWF AOT values are higher in the summer than in the winter, but the retrieved AOT tends to be zero in the winter while this is not the case in the ECWMF AOT product. We note that AOT can vary quickly and the exact match between the retrieved values and the ECWMF product is difficult because of the possible temporal and spatial mismatch of the two sets of values. A quantitative comparison between the retrieved AOT and ECMWF AOT data shows a low correlation with $R^2 < 0.3$ (results not shown). Nevertheless, the results indicate that it is possible to retrieve reasonable AOT values together with vegetation properties.

4. Discussion

4.1. Temporal autocorrelation to reduce the ill-posedness

While the meteorological scientific community has made substantial progress in the assimilation of satellite data and ground observations in time series for weather prediction purposes (Reichle 2008; Rodell et al. 2004), the use of prior information in land surface parameter retrieval is still at an early stage. Progress has been made to improve retrievals of vegetation properties (e.g., Verrelst et al. 2015; Wang et al. 2019), but the common approach is still to use single images independently.

Our approach aimed at improving the retrieval accuracy by reducing the ill-posedness, starting from the notion that the accuracy of retrieval
largely depends on the number of free parameters and the available observations of a target (Kimes et al. 2000). We found more stable retrievals by reducing this freedom by exploiting temporal autocorrelation, which can be viewed as either the increase in the number of the observations or the decrease in the freedom of the parameters. The relaxation times that we used to quantify the temporal autocorrelation were based on intuitive process understanding, but there is an indirect observational evidence in support. The soil and vegetation properties are responsible for the spectral shape of reflectance, and consequently, temporal autocorrelation of these properties will result in temporal autocorrelation of observed spectra. We examined the correlation of OLCI spectral observations acquired among pairs of days in an annual time series at the four study sites and found that the correlation coefficients generally decrease with increasing time interval (Fig. 11). On the one hand, the occurrence of positive correlations can be attributed to the temporal autocorrelation of land surface and atmospheric properties of a target. On the other hand, gradual changes of land surface and atmospheric properties and differences in sun-observer geometry lead to differences among the spectral observations. The spectral correlation coefficient initially decreases with time interval, but increases again when the time interval exceeds 180 days, thus reflecting a typical periodicity in natural processes, notably the annual cycle. The spectral observations on DOY 1 are highly correlated with the observations on DOY 365 since the temporal gap is 1 day if the seasonal cycle is considered. In this respect, it makes sense to work with the cosine of a phase angle rather than time intervals to quantify autocorrelation of time series, e.g., $\cos(\pi \times \text{DOY}/365)$, as a measure of temporal distance between observations.

In this study, only the previous observations are used to constrain the retrieval from successive observations. However, this works vice versa and the successive observations can also be used to constrain the retrieval from their predecessors. This could be achieved with iterations by first retrieving in a temporal forward and then background direction. In this respect, the proposed system is also a learning system, which should be able to improve its performance over iterations. The relaxation times we used in the time series retrieval imply a

---

**Fig. 8.** Retrieved values of LIDFa by using the standard retrieval approach (STDalgo) and the time-series retrieval approach (TSalgo) at four study sites. The yellow dots indicate when the retrieved LAI from OLCI is less than 0.5. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

**Fig. 9.** Retrieved values of leaf chlorophyll content $C_{ab}$ by using the standard retrieval approach (STDalgo) and the time-series retrieval approach (TSalgo) at four study sites. The yellow dots indicate when the retrieved LAI from OLCI is less than 0.5. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)
certain plasticity of parameters. Lauvernet et al. (2008) assumed invariant vegetation properties in their analysis of multi-temporal remote sensing data. This assumption reduces the ill-posedness by including more observations of the same (identical) target rather than incorporating the temporal autocorrelation in the model inversion problem. It also implies that the variation in spectral observations over time is solely explained by the change of sun-observer geometry. This constraint may work for observations closely separated in time (sub-daily), but it may result in large residuals in fitting the measured spectra if observations are separated by longer time intervals. A more plausible way is to allow for a moderate variation of vegetation properties over time. The relaxation times should strongly depend on the kind of land use, and on various geographical and climatological conditions. Our view is that an operational system should have sufficient flexibility to allow learning, so that relaxation times and other parameters can evolve with experience.

We categorize land surface and atmospheric properties into three groups according to their temporal variability. With the parameterization of SPART, the soil spectral shape parameters \( \varphi \) and \( \lambda \), which are mainly determined by the composition of the soil, leaf internal structural parameter \( N \) and the atmospheric property \( U_{O_3} \) are considered as unlikely to change over time. Hence, a long relaxation time is attributed to these parameters. In contrast, soil moisture, aerosol optical thickness and water vapour can change rapidly and thus, a short relaxation time is applied. The remaining model parameters, which include leaf pigment pool and canopy LAI, are assumed to have medium relaxation time. The relaxation time for LAI, \( LIDF_a \) and \( C_{ab} \) of 30 days is similar to that in Mousivand et al., (2015). However, the exact values used for relaxation time are semi-empirical and might change based on vegetation type and

![Fig. 10. Retrieved values of aerosol optical thickness (AOT_{500}) by using the standard retrieval approach (STDalgo) and the time-series retrieval approach (TSalgo) and ECWMF AOT data at four study sites.](image1)

![Fig. 11. Temporal autocorrelation of the TOA spectral observations of four study sites. Shown are the average values of the Pearson correlation coefficient of any two spectral observations for the same site in 2018 changing with the temporal gap between the two observations.](image2)
location. In practice, the values of the relaxation time can be established according to the rule of balancing the radiometric term and the prior information term in the cost function.

The retrieval algorithm uses a Bayesian approach to balance observations and prior knowledge, where ‘prior knowledge’ is obtained from observations at different times. This approach leads to much smoother time series of satellite derived land surface products than currently available, where real changes in land surface properties can be differentiated from noise (Figs. 3–5 and 8–10). The accuracy of the LAI estimation is improved by incorporating the temporal continuity of the land surface and atmospheric properties (Fig. 6).

4.2. Retrieval from TOA and TOC observations

Most of the retrieval algorithms solve a model inversion problem separately for the atmosphere and the surface in the following steps. First, atmospheric optical properties (e.g., transmittance) are estimated from measurements or products of atmospheric properties (e.g., AOT) and an atmospheric RTM. In order to perform this estimation, certain land surface optical properties should be assumed, e.g., zero reflectance of dark water bodies. Second, TOC reflectance is estimated from TOA reflectance by using the atmospheric optical properties, which is the so-called atmospheric correction. Third, canopy parameters are estimated by inverting a canopy RTM. This three-step procedure is, probably, sub-optimal, since surface and atmospheric variables are obviously strongly coupled, implying that knowledge of the characteristics of one component is required to estimate radiative properties of the other (Laurent et al. 2014; Yang et al. 2020a). Furthermore, accurate measurements of atmospheric properties are required to perform the atmospheric correction. Therefore, at least from this heuristic argument, using coupled surface and atmosphere RTMs is promising.

Several studies have coupled canopy RTMs with MODTRAN (Mousivand et al. 2015; Verhoef et al. 2018; Verhoef and Bach 2012, 2007), but the MODTRAN model is too heavy to invert. Many studies built LUTs or emulators to replace the complex atmosphere RTM. In the LUT approach, a discrete sample of model input is extracted from the full parameter space, and the corresponding model output is simulated by the forward radiative transfer model. As a result, the relationship between the sampled model input and output is straightforward. Similar to the LUT approach, emulators seek for simple relationships between model input and output usually relying on machine learning techniques (B&ger et al. 2020; De Grave et al. 2020; Estévez et al. 2020; Verrelst et al. 2019). The accuracy of LUT approaches and emulators is still debatable, especially for high-dimensional data, and the LUT and machine learning set-up is usually strongly determined by subjective decisions and preferences of the particular investigators and the kind of data. As an alternative solution, many studies directly treat the atmospheric effect on TOA reflectance as a prior information. Measurements of atmospheric properties are used to determine the transmittance and reflectance properties of the atmosphere layer and then upscale TOC to TOA observations. In this way, only canopy RTMs are inverted while the forward simulation of the MODTRAN model is needed. Compared to the atmosphere correction approach described above, the coupling model is theoretically better since it considers the radiative interactions between the surface and the atmosphere.

By using the coupled SPART model for inversion of the combined soil-vegetation-atmosphere system, we have demonstrated the possibility of retrieving vegetation properties, together with atmospheric properties, directly from TOA observations without atmospheric correction (Fig. 10). Compared to retrieval from TOC measurements, the retrieval from TOA measurements results in a higher number of unknown parameters. However, possible measurements of atmospheric properties can easily be incorporated as prior to constrain the model inversion problem. Hence, the combined surface-atmosphere retrieval is not necessarily more ill-posed than the retrieval from TOC data obtained after atmospheric correction of TOA measurements.

4.3. Multi-sensor retrieval of vegetation and atmospheric properties

Most remote sensing products are retrieved from single missions and sensors. The quality of land surface products is determined by the degree to which optical properties of the atmosphere and the surface can be differentiated during the overpass of a specific satellite (Xiao et al. 2017). The scope and limitations of each product are thus determined by the satellite (orbit, geometry and revisit time), the sensor and the retrieval algorithm (Garzón et al. 2008; Laurent et al. 2014). A challenge in quantitative remote sensing is not only to assimilate time series of satellite observations from one sensor (as we did in the present study), but also to assimilate observations from various other sensors. Differences in orbit, swath, and spectral configuration cause the viewing-observation geometry, spatial, spectral and temporal resolutions to vary from one sensor to the other. Although it is possible to combine higher level data products from these missions (e.g., interpreted data, such as leaf area index) to constrain land surface or atmospheric models through data assimilation schemes, differences in the retrieval algorithms used to produce the higher level data products limit their quantitative comparability in physical units.

The assimilation of spectral observations from various sensors can also be achieved by using the framework presented in this study, thanks to the coupled surface-atmosphere RTM. SPART takes sensor characteristics and observational geometries quantitatively into account in the radiative transfer modelling, which makes observations from different sensors and observational conditions compatible. Many sensors on satellite platforms measure radiances passively in the solar reflective part of the electromagnetic spectrum. For example, the missions Sentinel-2, 3, 4 and 5 (P) of the ESA’s Copernicus program all have sensors that measure reflected solar radiation in the visible and near-infrared (VNIR) part of the electromagnetic spectrum, but the data are used for different purposes (B&ger et al. 2012). The present retrieval framework offers a prototype for further development of a physically based multi-sensor approach for land-atmosphere interaction where the integration does not take place at the higher-level data products (e.g., LAI and fPAR), but at the lower-level product of TOA radiance. The physically based approach complements ongoing data driven initiatives (i.e. machine learning) towards multi-sensor data fusion (Estévez et al. 2020; Verrelst et al. 2015). Our proposed approach relies on a unified radiative transfer scheme for soil, vegetation, and the atmosphere (i.e., SPART). With such a scheme, the signal at TOA in a wide part of the electromagnetic spectrum is simulated as a function of the sun-observer geometry and the optical properties of the land surface and the atmosphere (e.g., Fig. 1). It is obvious that the ill-posedness of a model inversion problem reduces with an increasing number of observations of the same target. Therefore, time series of observations from multiple sensors and orbits can surely reduce the problem of ill-posedness in the retrieval process. Consequently, time-series observations from multiple sensors and orbits can be combined relatively easily.

4.4. Limitations and prospects

A new feature which has not been tried yet in the retrievals from satellite observations is the detection of human interferences on land surface. The model update rule was designed for conditions when the surface changes are mostly small and gradual due to natural processes. In that case, local linearization of the RTM can be applied, and a few updating steps (i.e., iterations) will be sufficient in most cases. On the other hand, if large discrepancies are found between predicted and observed satellite data, which cannot be attributed to clouds or soil moisture fluctuations, then these might be due to unexpected events such as clearings in a forest or harvesting activities in agriculture. Therefore, an “alarm” function may be included for special applications addressing this kind of events, which have an interest by themselves. In this way, detection of deforestation would form a natural by-product of the data assimilation system. The same would hold for cloud detection,
which would take place automatically, since the presence of clouds would mean such a strong contrast with a gradually changing land surface, that they could be detected very easily.

We examined the algorithms at a few sites, but the computational efficiency should be considered for global applications. Numerical optimization applied in this study is usually slower than LUT, machine learning or data derived approaches in terms of global applications, although TSalgo requires about 20% time compared with STDalgo thanks to a reasonable prior from the previous moments. If the dimensionality can be reduced by removing noise without losing information, numerical optimization might still be fast enough. On the other hand, remote sensing data are very suitable to parallelization. Hence, in practical situations an approach that proved to be successful in the lab, can be operationalized effectively, as has been demonstrated in the field of numerical weather prediction for instance. Alternatively, LUT or machine learning approaches can be employed, but adaption and adjustment of TSalgo for these approaches are required. Additionally, LUT and machine learning approaches can provide uncertainties of retrieved biophysical variables, which is not addressed in this study (Berger et al. 2020; De Grave et al. 2020; Estévez et al. 2020; Laurent et al. 2014).

In this study we addressed the well-known problem that the retrieval of land surface properties from either TOC and TOA spectral observations by inverting an RTM is ill-posed (Combal et al. 2003; Jacquemoud et al. 2009; Kimes et al. 2000). To reduce the underdetermined nature of the inversion problem, four types of regularization methods have been proposed in the literature (Baret and Buis 2008; Laurent et al. 2014) namely 1) using coupled RTMs, 2) incorporating a priori information, 3) employing spatial constraints and 4) employing temporal constraints.

Land surface observations by optical sensors include variations in the spectral, temporal and spatial domains, and directional variation in the radiometric data due to the anisotropy of land surface reflection. These four ‘dimensions’ of remote sensing data can be exploited, and a multi-sensor approach enriches the available information in these four domains. The most common way is to assume no variation of some land surface properties in the given field. For example, Atzberger et al. (2004, 2012) assumed leaf angle distributions of vegetation in an agricultural field to be the same, and Laurent et al. (2014) assumed LAI and leaf chlorophyll content of vegetation in a relatively homogenous field to be the same. These studies have shown the potential of using spatial constraints to improve the retrieval from mono-temporal observations. The present study focuses on the incorporation temporal autocorrelation as a priori information about land and atmospheric properties and investigates the importance of using this temporal constraint. We have demonstrated that our algorithm can be used to incorporate the temporal continuity of surface and atmospheric properties and also shown that the use of time series of observations can improve the quality of LAI retrievals according to the comparison to the MODIS LAI, field measured LAI and LAI estimates by using the standard retrieval algorithm (Figs. 4–6). Directional variation of remote sensing observations can be exploited by using RTMs, and the algorithm proposed in this study can be further adapted to include the spatial constrains or spatial autocorrelation, which will allow some degrees of spatial variation in a given field.

5. Conclusions

Time series of spectral observations from various satellite sensors allow a better monitoring of land surface properties. We present a retrieval framework that incorporates the temporal autocorrelation and the dependence of land surface and atmospheric properties to better use time series spectral observations. By using an invertible soil-vegetation-atmosphere radiative transfer model (i.e., the SPART model), we are able to retrieve land surface properties directly from OLCI TOA observations without atmospheric correction. We show that the temporal continuity of the land surface and atmospheric properties is a piece of promising prior information to reduce the ill-posedness of model inversion problems and improve the retrieval accuracy. In particular, it helps to mitigate unrealistic short-term changes in the retrieved variables.

The approach presented in this study can also be adapted to include the spatial autocorrelation and multi-sensor and multi-angle spectral observations in model inversion problems. The multi-sensor and multi-angle observations enrich the radiometric information, while the spatial and temporal autocorrelation of the land surface and atmospheric properties serve as prior information. This offers a way to maximize utilization of numerous remote sensing data and incorporate the available prior information either from measurements or from the spatial and temporal autocorrelations of the land surface and atmospheric properties. The prospect is to use such a system for a better quantitative use of remote sensing observations and a higher quality of the end products.

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.

Acknowledgements

The first author (Peiqi Yang) was supported by the Netherlands Organization for Scientific Research, grant ALWGO.2017.018. Egor Prikaziuk has received funding from the European Union’s Horizon 2020 research and innovation programme under the Marie Sklodowska-Curie grant agreement No 721995.

Appendix A

Fig. A1. Root-mean-square errors (RMSE) between modelled and measured TOA reflectance after optimization for the synthetic dataset.
Fig. A2. Root-mean-square errors (RMSE) between modelled and measured TOA reflectance after optimization for the real satellite dataset.

References

Atzberger, C., 2004. Object-Based Retrieval of Biophysical Canopy Variables Using Artificial Neural Nets and Radiative Transfer Models, 93, pp. 53–67. https://doi.org/10.1016/j.rse.2004.06.016.

Atzberger, C., Richter, K., 2012. Remote sensing of environment 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.

Baldocchi, D., Falge, E., Gu, L., Olson, R., Hollinger, D., Running, S., Anthoni, P., Bernhofer, C., Davis, K., Evans, R., 2001. FLUXNET: a new tool to study the temporal and spatial variability of ecosystem-scale carbon dioxide, water vapor, and energy flux densities. Bull. Am. Meteorol. Soc. 82, 2415–2434.

Baret, F., Buis, S., 2008. Estimating canopy characteristics from remote sensing observations: review of methods and associated problems. In: Advances in Land Remote Sensing: System, Modeling, Inversion and Application. Springer, pp. 173–201. https://doi.org/10.1007/978-1-4020-6450-0_7.

Baret, F., Guyot, G., 1991. Potentials and limits of vegetation indices for LAI and APAR assessment. Remote Sens. Environ. 35, 161–172. https://doi.org/10.1016/0034-4257(91)90009-U.

Berger, M., Moreno, J., Johannessen, J.A., Levelt, P.F., Hansen, R.F., 2012. ESA’s sentinel missions in support of earth system science. Remote Sens. Environ. 120, 84–90.

Berger, K., Atzberger, C., Danner, M., Wocher, M., Mauser, W., Hank, T., 2018. Model-based optimization of spectral sampling for the retrieval of crop variables with the PROSAIL model. Remote Sens. 10. https://doi.org/10.3390/rs10122063.

Berger, K., Verrelt, J., Ferré, J.-B., Hank, T., Wocher, M., Mauser, W., Camps-Valls, G., 2020. Retrieval of aboveground crop nitrogen content with a hybrid machine learning method. Int. J. Appl. Earth Obs. Geoinf. 92, 102174. https://doi.org/10.1016/j.jag.2020.102174.

Berger, K., Atzberger, C., Danner, M., Wocher, M., Mauser, W., Hank, T., 2018. Model-based optimization of spectral sampling for the retrieval of crop variables with the PROSAIL model. Remote Sens. 10. https://doi.org/10.3390/rs10122063.
Wang, D., Sagan, V., Guillevic, P.C., 2019. Quantitative remote sensing of land surface variables: progress and perspective. Remote Sens. 11, 11–15. https://doi.org/10.3390/rs11182150.

Weiss, M., Baret, F., Myneni, R.B., Pragnère, A., Knyazikhin, Y., 2000. Investigation of a model inversion technique to estimate canopy biophysical variables from spectral and directional reflectance data. Agronomie 20, 3–22. https://doi.org/10.1051/agro:2000105.

Widlowski, J.J., Pinty, B., Lavergne, T., Verstraete, M.M., Gobron, N., 2005. Using 1-D models to interpret the reflectance anisotropy of 3-D canopy targets: issues and caveats. IEEE Trans. Geosci. Remote Sens. 43, 2008–2017. https://doi.org/10.1109/TGRS.2005.859718.

Xiao, Z., Liang, S., Wang, J., Jiang, B., Li, X., 2011. Real-time retrieval of Leaf Area Index from MODIS time series data. Remote Sens. Environ. 115, 97–106.

Xiao, Z., Liang, S., Wang, J., Chen, P., Yin, X., Zhang, L., Song, J., 2014. Use of general regression neural networks for generating the GLASS leaf area index product from time-series MODIS surface reflectance. IEEE Trans. Geosci. Remote Sens. 52, 209–223.

Xiao, Z., Liang, S., Jiang, B., 2017. Evaluation of four long time-series global leaf area index products. Agric. For. Meteorol. 246, 218–230.

Xu, X., Du, H., Zhou, G., Li, P., 2016. Method for improvement of MODIS leaf area index products based on pixel-to-pixel correlations. Eur. J. Rem. Sens. 49, 57–72. https://doi.org/10.5721/EurRS20164904.

Yan, P., Chen, Y., Dadouma, A., Tao, Z., Sui, P., 2017. Effect of nitrogen regimes on narrowing the magnitude of maize yield penalty caused by high temperature stress in North China Plain. Plant Soil Environ. 63, 131–138. https://doi.org/10.17221/6/2017-PSE.

Yang, H., Yang, X., Heskel, M., Sun, S., Tang, J., 2017. Seasonal variations of leaf and canopy properties tracked by ground-based NDVI imagery in a temperate forest. Sci. Rep. 7, 1–10. https://doi.org/10.1038/s41598-017-01266-y.

Yang, P., van der Tol, C., Verhoef, W., Damm, A., Schickling, A., Kraska, T., Muller, O., Rascher, U., 2019. Using reflectance to explain vegetation biochemical and structural effects on sun-induced chlorophyll fluorescence. Remote Sens. Environ. 231. https://doi.org/10.1016/j.rse.2018.11.036.

Yang, P., van der Tol, C., Yin, T., Verhoef, W., 2020a. The SPART model: a soil-plant-atmosphere radiative transfer model for satellite measurements in the solar spectrum. Remote Sens. Environ. 247, 111870. https://doi.org/10.1016/j.rse.2020.111870.

Yang, P., Verhoef, W., van der Tol, C., 2020b. Unified four-stream radiative transfer theory in the optical-thermal domain with consideration of fluorescence for multi-layer vegetation canopies. Remote Sens. 12, 1–19. https://doi.org/10.3390/rs12233914.

Zarco-Tejada, P.J., Miller, J.R., Harron, J., Hu, B., Noland, T.L., Goel, N., Mohammed, G. H., Sampson, P., 2004. Needle chlorophyll content estimation through model inversion using hyperspectral data from boreal conifer forest canopies. Remote Sens. Environ. 89, 189–199. https://doi.org/10.1016/j.rse.2002.06.002.