ISPRS Journal of Photogrammetry and Remote Sensing 195 (2023) 408–417

Contents lists available at ScienceDirect

ISPRS Journal of Photogrammetry and Remote Sensing

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

# Estimation of vegetation traits with kernel NDVI

Qiang Wang [a,*], Álvaro Moreno-Martínez [a], Jordi Muñoz-Marí [a], Manuel Campos-Taberner [b], Gustau Camps-Valls [a]

[a] Image Processing Laboratory (IPL), Universitat de València, Catedrático A. Escardino, 46980, Paterna, València, Spain
[b] Earth sciences department, Universitat de València, C/ Doctor Moliner, 50, 46100, Burjassot, València, Spain

ARTICLE INFO

*Keywords:*
Vegetation index
NDVI
Kernel methods
Uncertainty quantification
Biophysical parameter estimation
Gross primary production
fAPAR
Leaf area index
Latent heat

ABSTRACT

Vegetation indices computed from spectral signatures are vastly used for monitoring the terrestrial biosphere. Indices are convenient proxies for canopy structure, and leaf pigment content, and consequently to estimate the photosynthetic activity of vegetation. Owing to its simplicity, the celebrated Normalized Difference Vegetation Index (NDVI) has been used as a proxy for greenness and canopy structure. Unfortunately, NDVI can only capture linear relationships of the near infrared (NIR) - red difference with the parameter of interest. To account for higher-order relations between the spectral channels, kernel NDVI (kNDVI) was proposed in (Camps-Valls et al., 2021). In this work, we give useful prescriptions for its proper use and show its good performance in a wider set of applications. We discuss the good characteristics of the index like boundedness, low error propagation. Furthermore, we give empirical evidence of performance in estimating in-situ vegetation parameters (leaf area index (LAI), gross primary productivity (GPP), leaf, and canopy chlorophyll content, green and total LAI and fraction of absorbed photosynthetically active radiation (fAPAR)) as well as the estimation of latent heat at flux tower level. We confirm the generally good performance of the index (correlation coefficient of kNDVI and canopy chlorophyll content is 0.919 and 0.933 for maize over two sites, as well as the correlation coefficient between kNDVI and carotenoid, is 0.816, 0.520 and 0.579 for three forest sites) and highlight its convenience in monitoring terrestrial ecosystems. To foster wider adoption of the new family of the index, we provide source code in 6 programming languages as well as efficient implementations in the Google Earth Engine (GEE) platform at https://github.com/IPL-UV/kNDVI.

## 1. Introduction

Monitoring vegetation conditions and functioning with remote sensing is key to a wide range of disciplines, from precision agriculture to global carbon cycle research. Earth observation through satellite remote sensing allows monitoring of terrestrial vegetation status, phenology, and health globally, by deriving spatially explicit and temporally resolved products. Quantifying vegetation cover, the biochemistry and structure through remote sensing means is paramount to understand pressing global climate change as well as to work towards enhancing biodiversity and improving agriculture productivity. With the development of remote sensing, sensors onboard airborne and satellites provide a convenient way to reveal the vegetation status. Researchers have developed a vast number of vegetation indices (VIs) based on the mathematical operation of the spectral reflectance in different bands, which can maximize the sensitivity to particular biophysical phenomena (e.g., greenness or photosynthetic activity). Vegetation indices are spectral transformations of two or more bands designed to enhance the contribution of vegetation properties and allow reliable spatial and temporal inter-comparisons of terrestrial photosynthetic activity and canopy structural variations. They allow us to monitor seasonal, inter-annual, and long-term variations of vegetation structural, phenological, and biophysical parameters. Therefore, spectral indices can be considered as meaningful extracted features that help to estimate the biophysical variables. Many narrow-band vegetation indices have been proposed (Xue and Su, 2017), mainly from empirical approaches (Kauth and Thomas, 1976; Daughtry et al., 2000; Haboudane et al., 2004) and with parameterization obtained by statistical modeling. The vast majority of remote sensing publications and applications still rely on two-band difference ratios, and in particular, on the normalized difference vegetation index (NDVI) (Rouse Jr. et al., 1974; Tucker, 1979), most probably because of ease of use and existing long-time records. The NDVI exploits the fact that green healthy vegetation shows contrasting behavior in how it reflects red and near-infrared (NIR) radiation. The more chlorophyll there is in a canopy (and thus the higher the

* Corresponding author.
*E-mail addresses:* qiang.wang@uv.es (Q. Wang), alvaro.moreno@uv.es (Á. Moreno-Martínez), jordi.munoz@uv.es (J. Muñoz-Marí), Manuel.Campos@uv.es (M. Campos-Taberner), gustau.camps@uv.es (G. Camps-Valls).

https://doi.org/10.1016/j.isprsjprs.2022.12.019
Received 16 July 2022; Received in revised form 21 December 2022; Accepted 22 December 2022
Available online 27 December 2022

photosynthetic potential), the more the canopy absorbs red radiation, whereas the presence of more biomass will scatter (and thus reflect) more NIR radiation, useless for photosynthesis. This simple mechanism explains the definition of NDVI, which essentially calculates the difference between red and NIR reflectances. NDVI has been very effective in assessing chlorophyll and other pigments content (Myneni et al., 1995; Haboudane et al., 2004), and acting as an excellent proxy of vegetation density parameters, like the leaf area index (LAI) and the fractional vegetation cover (FVC) (Le Maire et al., 2008; Berni et al., 2009).

However, NDVI may not relate linearly with biophysical parameters of interest, like green biomass. Well-known issues of NDVI with nonlinear relations (Huete et al., 1997; Paruelo and Lauenroth, 1998; Zhu et al., 2020) and saturation effects for higher vegetation cover area and during the peak of the season (Sellers, 1985; Gao et al., 2000; Thenkabail et al., 2000) have been reported. Several attempts to correct such limitations have been recently proposed that consider empirical transformations of the index, such as using exponentiations like $NDVI^2$ (Wang et al., 2017) or more general (yet arbitrary) ones like in Wu (2014). Motivated from a more physiological standpoint and using radiative transfer models (Sellers, 1985), some authors have proposed to multiply NDVI with NIR to calculate a new index called NIRv (Badgley et al., 2017). Unfortunately, despite its good performance, NIRv is not a dimensionless quantity (or actual index), and it is uncertain how it can cope with nonlinearities and saturation effects.

In Camps-Valls et al. (2021), we introduced a methodology to make vegetation indices nonlinear (Rojo-Álvarez et al., 2018; Camps-Valls and Bruzzone, 2009). The kernel NDVI (kNDVI) recovers all higher-order differences between the NIR and red reflectance bands, and when an appropriate kernel function is used, the index comes in a rather simple and practical expression. The index demonstrated to be a better proxy to estimate GPP at flux tower level, SIF from GOME-2, and LAI from MODIS. Other authors have successfully used kNDVI to estimate GPP at high spatial resolution using Sentinel-2 (Pabon-Moreno et al., 2022), in the characterization of light-use efficiency models (Pei et al., 2022), to better explain CO2 fertilization effects on vegetation photosynthesis (Wang et al., 2021), and to identify drivers of decline in temperate deciduous forests (Hase et al., 2022), among many other applications. However, other authors have struggled with the use of the index and obtained poor results for the relationship between kNDVI and vegetation traits, for instance, GPP (Wang et al., 2022).

In this paper, we argue that these contrasting behavior can be explained by the fact that kNDVI has one or more parameters (depending on the used kernel) to be adjusted or estimated from the observations, which requires some attention. In Section 2, we review the theory and properties of kNDVI and clarify the correct use of the index in a set of prescriptions and practical properties (bounded index, low uncertainty propagation). Section 3 shows results in estimating in-situ vegetation parameters (LAI, GPP, leaf, and canopy chlorophyll content, green and total LAI, green and total fAPAR) as well as the estimation of latent heat at flux tower level. Some concluding remarks are given in Section 4.

## 2. Nonlinear kernel vegetation indices

This section reviews the nonlinear NDVI introduced by Camps-Valls et al. (2021), paying particular attention to their theoretical properties and providing some relevant prescriptions for the proper use and application to clarify some misconceptions and misuse in the recent literature (Gensheimer et al., 2021; Wang et al., 2022). Even in recent reviews on vegetation indices, the oversimplification of kNDVI continues (Zeng et al., 2022).

### 2.1. Nonlinear vegetation indices with kernels

The use of kernel methods to make NDVI nonlinear in Camps-Valls et al. (2021) was not incidental, as they do precisely what otherwise is done empirically: learn a nonlinear function to be applied to the involved bands to retrieve the maximum information possible. Kernel methods map the involved data instances (here, the involved spectral bands) using a nonlinear feature map to a higher-dimensional (possibly infinite) Hilbert feature space where the operation (the NDVI) index is defined. Intuitively, one can see kernel methods as a principled way to *linearize* any problem, which is what some vegetation indices are designed to do empirically.

### 2.2. Kernel NDVI

Deriving nonlinear (kernel) indices requires the definition of a feature mapping $\phi(\cdot)$ to a Hilbert space $\mathcal{H}$ endorsed with the kernel reproducing property (Camps-Valls and Bruzzone, 2009; Rojo-Álvarez et al., 2018; Schölkopf and Smola, 2002). Let us exemplify the use of kernel methods to derive difference ratio indices involving only two bands, such as in the popular NDVI. We focus here on the kernel NDVI, but other kernel indices could be derived (Camps-Valls et al., 2021). The normalized difference vegetation index is defined as $\text{NDVI} = \frac{n-r}{n+r}$, where $n$ and $r$ are the reflectances in the NIR and the red bands, respectively. For deriving the kNDVI, we define two scalars $n$ and $r$ in the real domain, and a feature map $\phi \mapsto \phi(n) \in \mathcal{H}$ with an associated reproducing kernel $k(n,\cdot) = \langle \phi(n), \cdot \rangle_{\mathcal{H}}$, and likewise for $r$. Now let us define two feature maps that work on the joint $z := (n, r)$ feature vector $\psi(z) := \phi(n) - \phi(r)$, $\varphi(z) := \phi(n) + \phi(r) \in \mathcal{H}$, with an associated difference and summation kernels (Camps-Valls et al., 2006):

$$m(z,z) = \langle \psi(z), \psi(z) \rangle_{\mathcal{H}} = k(n,n) + k(r,r) - k(n,r) - k(r,n) \quad (1)$$

$$\ell(z,z) = \langle \varphi(z), \varphi(z) \rangle_{\mathcal{H}} = k(n,n) + k(r,r) + k(n,r) + k(r,n) \quad (2)$$

The reproducing property allows us to estimate distances, similarities and angles between such mapped reflectances, and allows us to define the NDVI in Hilbert space via kernel functions:

$$\text{kNDVI} = \frac{m(z,z)}{\ell(z,z)} = \frac{k(n,n) + k(r,r) - k(n,r) - k(r,n)}{k(n,n) + k(r,r) + k(n,r) + k(r,n)} \quad (3)$$

Yet, the pertinent question is to ask about what is the actual form of kernel functions $k$. Different kernel functions can be used, which essentially implement an implicit feature map (Rojo-Álvarez et al., 2018). For example, a linear kernel function $k(a,b) = a\ b$ implements the identity map $\phi(a) = a$, while a $p$-order polynomial kernel function $k(a,b) = (1 + a\ b)^p$ maps the data points to a $p$ dimensional space with all $p$ monomials, $\phi(a) = [1, a, a^2, \ldots, a^p]$. The popular radial basis function (RBF) kernel, $k(a,b) = \exp(-\frac{1}{2\sigma^2}(a-b)^2)$, defines an implicit feature map to an infinite-dimensional space, with an appealing property as follows.

The Gaussian kernel exploits all higher-order relations between the considered reflectances. Let us assume the kernel $k(a,b) = \langle \phi(a), \phi(b) \rangle = \exp(-\gamma(a-b)^2)$, where $\gamma = 1/(2\sigma^2) \in \mathbb{R}^+$. Then, the explicit feature map $\phi$ is infinite dimensional:

$$\phi(a) = \exp(-\gamma a^2) \left[1, \sqrt{\frac{2\gamma}{1!}} a, \sqrt{\frac{(2\gamma)^2}{2!}} a^2, \sqrt{\frac{(2\gamma)^3}{3!}} a^3, \ldots\right]^\top \quad (4)$$

The kernel $k(n,r) = \langle \phi(n), \phi(r) \rangle_{\mathcal{H}}$ is therefore a dot product between infinite-dimensional expansions of $n$ and $r$. The RBF kernel function summarizes all higher-order differences between the NIR and red reflectance bands as:

$$k(n,r) = \sum_{t=0}^{\infty} (-1)^t \gamma^t (n-r)^{2t} / t! \quad (5)$$

and note that $\sigma$ has an important role in the definition of the kernel measure and thus the kNDVI.

The RBF is perhaps the most widely used kernel function in the literature because of this property and the practical advantage of having just one parameter to tune, the kernel lengthscale $\sigma$ (or $\gamma$) (Rojo-Álvarez et al., 2018). This was the kernel used in Camps-Valls et al. (2021) to derive a simplified form of kNDVI. Note that when using NIR, the RBF kernel function is implicitly mapping all monomials of the differences

$\phi(n,r) \propto \{n-r,(n-r)^2,(n-r)^3,\ldots\}$, and then kNDVI summarizes such higher orders into a scalar.

Now, we are ready to define the NDVI in Hilbert spaces by adopting the RBF kernel function, $k(\mathrm{n},\mathrm{r}) = \exp(-\frac{1}{2\sigma^2}(\mathrm{n\text{-}r})^2)$, where the $\sigma$ parameter controls the notion of distance between the NIR and red bands, $n$ and $r$. Note that the self-similarity $k(a,a) = k(b,b) = 1$ for the RBF kernel, and that by definition kernel functions are symmetric $k(a,b) = k(b,a)$, which allows us to simplify the kNDVI in (3):

$$\mathrm{kNDVI} = \tanh\left(\left(\frac{n-r}{2\sigma}\right)^2\right), \tag{6}$$

where $\sigma$ is a parameter that must be specified in each particular application. When a linear kernel is used one can demonstrate that the kNDVI gives the same result as its linear counterpart, the celebrated NDVI, while providing theoretical guarantees of performance.

### 2.3. Properties and prescriptions

#### 2.3.1. How to estimate an optimal $\sigma$

The crucial question remains: even if we use the RBF kernel function, what is the optimal value for the $\sigma$ parameter? Setting the kernel parameters is critical and has an essential impact on the solution (Rojo-Álvarez et al., 2018). Furthermore, this is even more challenging in unsupervised scenarios like the one we face here with the definition of indices.

A standard heuristic considers setting the lengthscale parameter $\sigma$ to the average value of the involved objects (NIR and red reflectances in our case), that is $\sigma = 0.5(n+r)$. However, this can lead to suboptimal results in skewed or undersampled data distributions: Note that $\sigma$ should reflect the notion of similarity between NIR and red reflectances of the whole application domain. Thus one should consider the spatial or temporal neighborhood. The 'mean heuristic' may be also cast as a rough estimation of the pixel's albedo (see discussion in supplementary materials of Camps-Valls et al., 2021). This heuristic also leads to an interesting (over)simplification of kNDVI:

$$\mathrm{kNDVI} = \tanh(\mathrm{NDVI}^2) \tag{7}$$

Two nonlinearities are used here; NDVI is squared similarly to Wang et al. (2017), and the result is squashed with a sigmoid function (to adapt to very high and sparsely vegetated regions intuitively). The double effect has shown a reduction in bias (Camps-Valls et al., 2021). However, using such a heuristic is an oversimplification of the kNDVI index. It is worth mentioning again that the formal way to use kNDVI requires first selecting the kernel function, which can be the RBF or any other one more suited to reflect similarities between NIR and red bands. In addition, the associated hyperparameter has to be adjusted, $\sigma$ for the RBF, but others for the wealth of kernel functions present in the literature. Probably because it is convenient, easy, and parameter-free, some authors blindly adopted this simplified kNDVI, which could explain some suboptimal results (Gensheimer et al., 2021; Zeng et al., 2022).

Let us indulge ourselves with some clarifications by assuming the adoption of the RBF kernel function. In this case, the parameter $\sigma$ directly affects the non-linearity and typically strongly impacts the index performance. Choosing a different $\sigma$ per biome or climatic region is certainly a good option, but results do not necessarily improve our experience. On the other hand, estimating the $\sigma$ parameter from the data involved makes a difference. Several options exist to estimate $\sigma$ from data:

1. Estimate $\sigma$ independently for each pixel as a proxy to pixel's albedo, $\sigma = 0.5(\mathrm{NIR+red})$, which oversimplifies the equation of kNDVI, $\mathrm{kNDVI} = \tanh(\mathrm{NDVI}^2)$, and typically leads to poor performance. Related heuristics have been recently introduced in Pabon-Moreno et al. (2022).
2. Estimate $\sigma$ from the region or biome by the average distance between all $N$ pixels therein, $\sigma = N^{-1}\sum_i^N |\mathrm{NIR}_i - \mathrm{red}_i|$; or
3. when working with (nonstationary) time series, one can estimate $\sigma$ from the temporal distances, $\sigma = T^{-1}\sum_{t=1}^{T} |\mathrm{NIR}_t - \mathrm{red}_t|$.

It is worth noting that using a $\sigma$ value per pixel (even if using a spatial or temporal mean heuristic) makes kNDVI adapt to completely different dynamic ranges. Such an approach is convenient for addressing challenging (extreme) cases of very densely and sparsely vegetated regions, as shown in Camps-Valls et al. (2021) to estimate GPP, LAI, and SIF.

#### 2.3.2. How to deal with radiances and reflectances

Working with radiances or reflectances changes the scale of the signal and thus of the range of reasonable $\sigma$ values, so we recommend either (a) scaling the data to reflectance values before fixing the $\sigma$ value or (b) estimating the $\sigma$ value from radiance data by the average distance criterion (see discussion above).

#### 2.3.3. Dealing with sensor noise, clouds, and water bodies

Vegetation indices should not be applied blindly over non-vegetated regions, such as cloudy pixels or water bodies. This issue is even more severe in kNDVI which depends explicitly on a parameter to be estimated from observations. Therefore, one should remove those cases from the calculation of the mean heuristic through time or space, or replace the mean with the median, which worked well in our case studies.

#### 2.3.4. The applicability of kNDVI

Note that the value of the kNDVI is bounded to $[0,1]$ by construction. It is recommended to apply kNDVI for vegetated pixels only, so it is advisable to mask water/snow pixels. Actually, in the Google Earth Engine (GEE) (Gorelick et al., 2017) demos in https://github.com/IPL-UV/kNDVI, we illustrate the correct use of kNDVI for spatial or temporal (phenological) studies: in all cases, we work with reflectance data and mask out clouds and water bodies when estimating a proper kernel length scale. Code in the other 7 programming languages can be obtained in the same repository.

#### 2.3.5. kNDVI sensitivity and uncertainty with changing $\sigma$

The index's sensitivity is governed by the selected $\sigma$ (Camps-Valls et al., 2021): The higher the $\sigma$ value, the lower the derivative of the kNDVI with respect to the NIR and red bands, and hence more sensitive to densely vegetated regions. On the contrary, the lower the $\sigma$ value, the more sensitive will be the kernel index to sparsely vegetated regions. This critical property can help select the parameter to increase the sensitivity in specific biomes or applications.

Nevertheless, also the derivatives allow us to study the propagation of uncertainty in the reflectance channels. Results in Camps-Valls et al. (2021) demonstrated theoretically and empirically that the kNDVI propagates a lower amount of error than the rest of the indices, like NDVI or $\mathrm{NIR}_v$. We will confirm this ability in a broad range of vegetation traits.

## 3. Experimental results

We illustrate the performance of the kNDVI in several remote sensing problems and compare it to indices in the literature using the same bands (NDVI and NIRv) involving: (1) crop monitoring of different biophysical parameters (GPP, Chl, LAI, PAR, and fAPAR), (2) estimation of leaf traits (including chlorophyll-a and chlorophyll-b, carotenoids, leaf mass per area, nitrogen -and carbon-concentration by mass) over forests, and (3) the estimation of latent heat at flux tower level.

### 3.1. Accurate proxy for vegetation parameters over cropland

#### 3.1.1. Study area and in-situ measurements

In this study, we investigated two AmeriFlux sites at the University of Nebraska-Lincoln Agricultural Research and Development Center where continuous monitoring was applied during the growing season between 2001 and 2008. Both sites are equipped with irrigation systems. Maize is planted at site 1 continuously while crop rotation is implemented on site 2 with maize planted in odd years and soybean grown in even years. Detailed information regarding the observatory can be found in Peng et al. (2011).

For the two sites, several vegetation parameters were recorded: including GPP, incident potential photosynthetically active irradiance (PARpot), incoming Photosynthetically Active Radiation (PARin), leaf Chlorophyll (leaf Chl), canopy Chlorophyll (canopy Chl), green leaf area index (green LAI), total leaf area index (total LAI), the fraction of radiation absorbed by photosynthetically active "green" vegetation (green fAPAR), the fraction of radiation absorbed by photosynthetically total active vegetation (total fAPAR) as well as soil moisture at 10 and 25 cm. Meanwhile, eddy covariance sensors were established to capture the hourly CO2 fluxes. PAR is measured with point quantum sensors (LI-190, LI-COR Inc, Lincoln, NE). LAI is measured via LI-3100 by averaging six samples from the sites and Chl is obtained by multiplying the LAI and the leaf Chl retrieved from the reflectance observation from Ocean Optics radiometer (details can be found in Gitelson et al., 2003). Meanwhile, hyperspectral radiometers are utilized to measure the canopy reflectance with a spectrum spanning 400 to 1000 nm with a 1.5 nm interval. The number of measurements recorded in 2001 and from 2006 to 2008 are limited and do not cover the crop growth period. Thus, only measurements during 2002 and 2005 were used for later analysis in this study.

#### 3.1.2. Seasonal variability of vegetation parameters

Fig. 1 shows the time series of observation of GPP and canopy Chl of the maize growth period (June 10th to September 20th) at site 1 between 2002 and 2005. The crop growth cycle can be easily identified with the GPP being very low (around 5 $gCm^{-2}d^{-1}$) at the beginning of the crop growth period. GPP starts to increase from June onwards and reaches its maximum on July 15th,2022, July 17th, 2023, July 19th, 2024, and July 12th, 2025 with the corresponding values are 27.45, 24.67, 30.41, and 24.95 $gCm^{-2}d^{-1}$. Then, the GPP declines until the middle or end of September. A similar trend is being found for canopy Chl. However, the peak value of canopy Chl was found in late July or early August, two weeks after GPP reached its peak value, showing the dependence of the GPP on the sun's illumination, which is strong in the middle of July for the northern hemisphere. The chlorophyll varied with the crop growth stage, which reaches the maximum at the beginning of fluorescence emergence, heading from Lancashire et al. (1991). Time series of other biophysical parameters ($GPP/PAR_{pot}$, $GPP/PAR_{in}$, green LAI, total LAI, green fAPAR, and total fAPAR can be found in Fig. S1. It can be seen that green LAI, total LAI, green fAPAR, and total fAPAR follow the same seasonal change of GPP. However, total LAI and total fAPAR show small variations compared with green LAI and green fAPAR during the later stage of maize growth. These differences could be attributed to decreased photosynthetic components during the reproductive stage and drastic drops in senescence, which causes the decrease of green LAI and fAPAR. However, the PAR absorption from non-photosynthetic components is relatively stable during the reproductive and senescence stages. Hereby, we plotted the ratio of carbon uptake to absorbed photosynthetically active radiation $GPP/PAR_{pot}$ and $GPP/PAR_{in}$, which is also represented by light use efficiency (LUE) shown in Fig. S1(a) and (b). It shows a similar seasonal change as chlorophyll, which increases from the beginning of the crop growth till late July or early August, then declines till the end of crop growth. Similar seasonal changes for site 2 can be found in Figs. S2 and S3 for the maize and soybean growth period, respectively.

**Table 1**
Correlation coefficient (absolute value) between various VIs and biophysical parameters for cropland ($p$-value is being indicated in the bracket).

| Site | Crop type | Parameter | NDVI | NIRv | kNDVI |
|---|---|---|---|---|---|
| 1 | Maize | GPP | 0.711 (0) | 0.790 (0) | **0.798** (0) |
| | | GPP/PARpot | 0.783 (0) | 0.813 (0) | **0.830** (0) |
| | | GPP/PARin | 0.804 (0) | 0.877 (0) | **0.884** (0) |
| | | Leaf Chl | 0.254 (0.016) | 0.466 (0) | **0.526** (0) |
| | | Canopy Chl | 0.820 (0) | 0.916 (0) | **0.919** (0) |
| | | green LAI | 0.925 (0) | 0.941 (0) | **0.948** (0) |
| | | total LAI | **0.933** (0) | 0.855 (0) | 0.885 (0) |
| | | green fAPAR | 0.948 (0) | 0.940 (0) | **0.954** (0) |
| | | total fAPAR | **0.919** (0) | 0.766 (0) | 0.844 (0) |
| | | median | 0.820 | 0.855 | **0.884** |
| 2 | Maize | GPP | 0.791 (0) | 0.869 (0) | **0.872** (0) |
| | | GPP/PARpot | 0.844 (0) | 0.893 (0) | **0.894** (0) |
| | | GPP/PARin | 0.849 (0) | 0.894 (0) | **0.911** (0) |
| | | Leaf Chl | 0.223 (0.178) | 0.366 (0.024) | **0.406** (0.012) |
| | | Canopy Chl | 0.861 (0) | 0.932 (0) | **0.933** (0) |
| | | green LAI | 0.938 (0) | 0.958 (0) | **0.963** (0) |
| | | total LAI | **0.933** (0) | 0.918 (0) | 0.921 (0) |
| | | green fAPAR | 0.943 (0) | 0.924 (0) | **0.955** (0) |
| | | total fAPAR | **0.906** (0) | 0.831 (0) | 0.872 (0) |
| | | median | 0.861 | 0.894 | **0.911** |
| | Soybean | GPP | 0.730 (0) | **0.747** (0) | 0.734 (0) |
| | | GPP/PARpot | 0.809 (0) | **0.820** (0) | 0.807 (0) |
| | | GPP/PARin | 0.817 (0) | **0.856** (0) | 0.850 (0) |
| | | Leaf Chl | 0.342 (0.018) | 0.422 (0.003) | **0.449** (0.002) |
| | | Canopy Chl | 0.750 (0) | 0.833 (0) | **0.841** (0) |
| | | green LAI | 0.891 (0) | **0.926** (0) | 0.918 (0) |
| | | total LAI | 0.893 (0) | **0.922** (0) | 0.917 (0) |
| | | green fAPAR | **0.949** (0) | 0.914 (0) | 0.920 (0) |
| | | total fAPAR | **0.976** (0) | 0.932 (0) | 0.954 (0) |
| | | median | 0.817 | **0.856** | 0.850 |

#### 3.1.3. Applicability of kNDVI

We investigated if the recently proposed kNDVI is viable to estimate crop parameters by using in-situ measurements. For this, the widely used RBF kernel mentioned before is being used with the $\sigma$ range of (0.01,10) in the logarithm scale. Then we calculated the correlation between the 1000 corresponding kNDVI and searched the optimal $\sigma$ based on the maximum linear correlation between the corresponding kNDVI and the specific vegetation parameter. We carried out the analysis individually for these two sites as well as the crop type. Then, optimal correlation coefficients which show a maximum linear correlation between various vegetation parameters and NDVI, NIRv, and kNDVI are shown in Table 1.

Table 1 shows that kNDVI yields the largest correlation with the most of vegetation parameters (GPP, $GPP/PAR_{pot}$, $GPP/PAR_{in}$, Leaf Chl, Canopy Chl, green LAI, green fAPAR) compared with NDVI and NIRv for site 1 (90 points). This implies that compared with NDVI and NIRv, kNDVI is more suitable to be used for estimating the photosynthetic activity (GPP, green LAI, green fAPAR), LUE ($GPP/PAR_{pot}$, $GPP/PAR_{in}$ as well as the Leaf Chlorophyll for maize with irrigation. However, NDVI shows a larger correlation with total LAI and total fAPAR than kNDVI. This difference is mostly because, on September 17, the reflectance of bands larger than 736 nm drops dramatically with a 0.1 decrease compared with the previous observation on September 13th. Consequently, kNDVI decreases drastically. This causes a high skewed kNDVI (skewness is −2.09) time series in the end, while the skewness of NDVI and original total fAPAR are −1.82 and −1.65, respectively. This indicates that if the original vegetation trait is highly skewed, kNDVI may not be able to show better performance compared with NDVI. Similar findings hold for site 2 over the maize growth period with 38 data points. To further justify this argument, a more complete dataset covering a whole crop growth period is required, which is not the case in our experiment.

For soybean in site 2 (47 data points), NIRv is mainly correlated with GPP and LAI, while kNDVI shows the highest correlation with

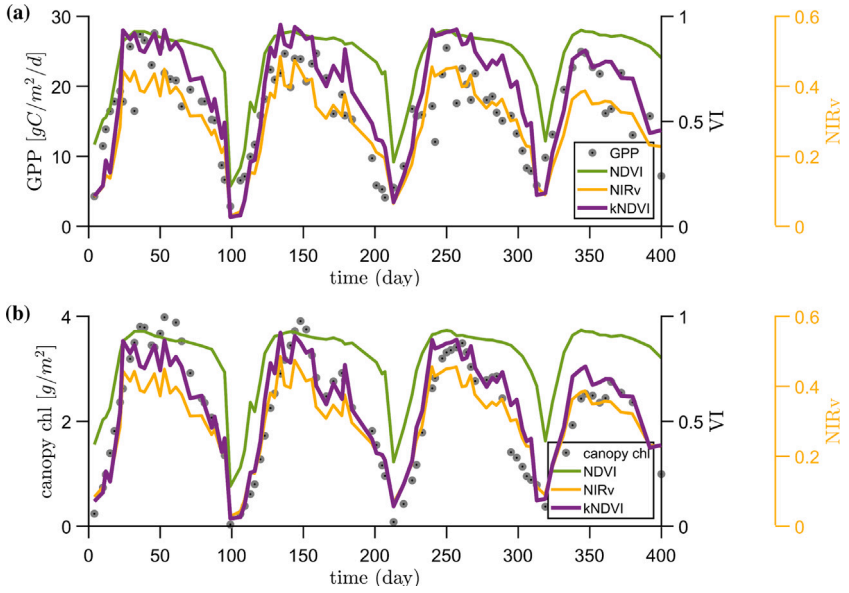


**Fig. 1.** Time series of GPP, canopy chl, and derived vegetation indices (NDVI, NIRv, kNDVI) for site 1.

chlorophyll and NDVI shows the best correlation with fAPAR. The possible reason for NIRv showing the best correlation with GPP is that the range of GPP variation for soybean is relatively small, which might be suitable for NIRv since NIRv shows a small dynamic change over the soybean growth. It should be noticed that kNDVI and NIRv perform similarly as proxies to GPP (differences in correlation around 0.01). It also should be noticed that for fAPAR, NDVI is the best proxy. The possible reason is that NIRv does not perform well is that the fAPAR from middle August to early September keeps stable at a value around 0.95, however, the NIRv decreases from 0.49 to 0.43 (see Fig. S3(h) and (i)), which is very high for its range during the whole crop cycle (0.05 – 0.58). Furthermore, regarding the temporal pattern of different indices, both NIRv and kNDVI reached peak value on August 5th, 2004 while NDVI reaches the peak value on August 20th, which is close to the time when fAPAR reaches the maximum (August 30th), as indicated in (i) of Fig. S3.

To illustrate further the correlation indicated in Table 1, the scatter plot between various VIs and GPP as well as canopy Chl are shown in Fig. 2. NDVI exhibits a clear saturation effect, indicated by large values in the range of 0.75–1, where a high non-linearity should be considered. Especially in Fig. 2(a), a cluster can be identified with NDVI and GPP in the range of 0.75–0.85 and 3–10 $gCm^{-2}d^{-1}$. Meanwhile, a quasi-linear relationship can be found with the NDVI in the range of 0.3–0.75. By contrast, both kNDVI and NIRv are less affected by the saturation problem, and advantageously kNDVI values are least deviated from their regression line compared with NIRv and NDVI.

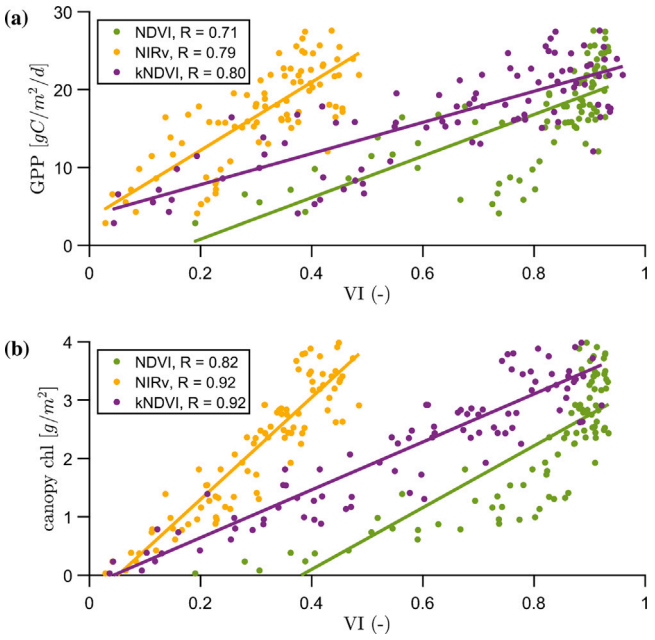


**Fig. 2.** Scatter plot for GPP (a) and canopy chl (b) against VIs (NDVI, NIRv, kNDVI) for maize over site 1.

### 3.1.4. The effect of kernel lengthscale parameter

Let us illustrate the RBF kernel lengthscale sensitivity using GPP and canopy chlorophyll. To do so, we plotted the correlation between kNDVI and these two variables with varying $\sigma$ values in the interval of $(0.01, 100)$ for the maize crop type at site 1. Fig. 3 indicates that with $\sigma$ increases from 0.01 onward, the correlation between GPP and kNDVI improves substantially and starts to outperform the NDVI. When sigma is very low, the $(n-r)/(2\sigma)$ is extremely large (see the right side of the black line shown in Fig. 4), and leads to a saturated kNDVI value of 1. On the contrary, as $\sigma$ increases, the saturation is less severe. Meanwhile, with a suitable $\sigma$, kNDVI can expand large values, minimize saturation effects, and shrink low values with the property of tanh and square. Under these circumstances, the kNDVI can better capture GPP temporal variability, and we achieve a significant correlation improvement. This can be seen in Fig. 3 with $\sigma$ values within the range of (0.01,0.20). In this example, when $\sigma$ equals 0.20, the functionality of tanh benefits the kNDVI, resulting in a high correlation of 0.80. With a further increase of $\sigma$, the $(n-r)/(2\sigma)$ inside the tanh operator for kNDVI decreases, consequently reducing the dynamic range of the kNDVI. When $\sigma$ is smaller than 0.2 (on the left side of blue line in Fig. 4, the shrinkage functionality of tanh is simplified to a linear function, and the correlation between kNDVI and GPP stays constant. Finally, for $\sigma$ values higher than 1.66, the correlation between kNDVI and GPP is stable at 0.76, and no further correlation improvement is noticeable.

It also can be noticed that the stable value of the correlation between kNDVI and GPP is slightly lower than the correlation between NIRv and GPP. This is because when the sigma increases, the range

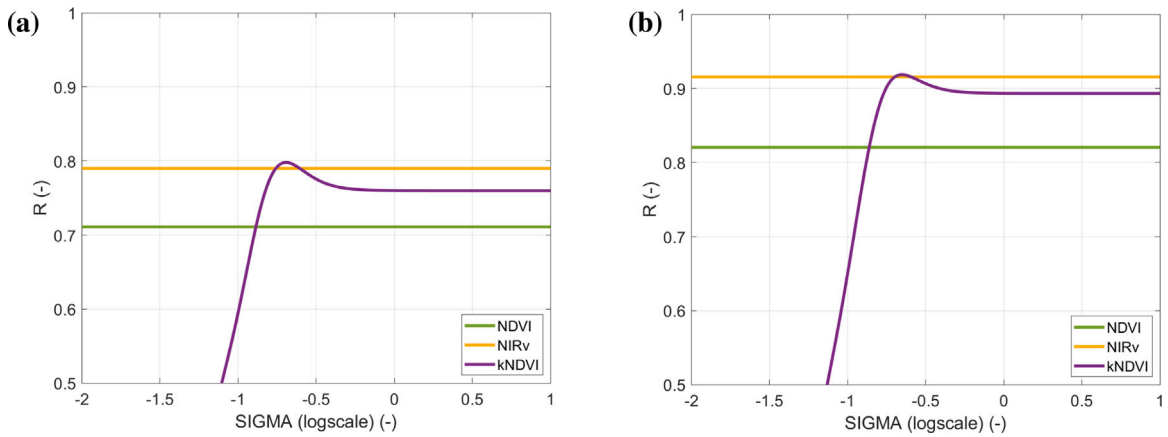


**Fig. 3.** Correlation of kNDVI and biophysical parameters with various sigma: GPP (a) and canopy chl (b) versus $\sigma$ for maize over site 1.

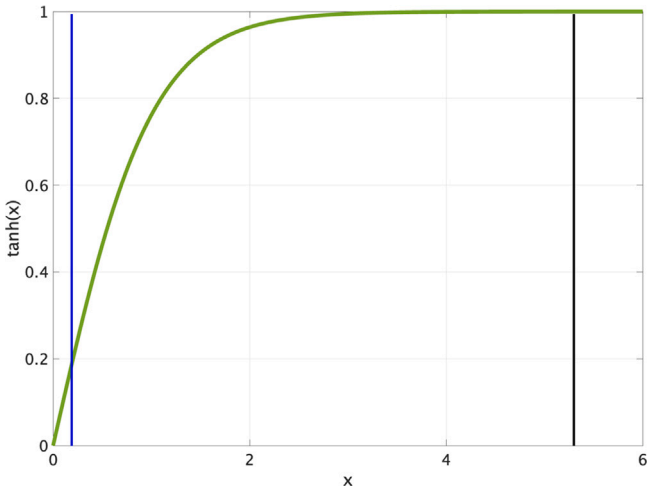


**Fig. 4.** The tanh function over the range (0, 6). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

of kNDVI shrinks according to Eq. (6), this results in a small range of kNDVI, which eventually makes the correlation between kNDVI and GPP smaller than the correlation between NIRv and GPP. For instance, when the sigma is equal to 0.5 (so −0.3 in log scale), the kNDVI will be between 0.0070 and 0.2993 while the NIRv ranges from 0.0290 to 0.4852. Secondly, there is a truncation error during the correlation analysis where we keep the correlation with 4 decimals (0.0001), this may cause the error of 0.00025. Even stable value of the correlation between kNDVI and GPP is lower than NIRV and GPP, this does not mean the correlation coefficient between kNDVI and vegetation traits will always stay at a level inferior to the correlation between NIRV and vegetation traits. To further support our argument, the correlation between various canopy chlorophyll and VIs or site 2 over the soybean growth period is shown in Fig. S4. Here, when we set the sigma as 0.5, the kNDVI ranges from 0.0108 to 0.4075, while the corresponding NIRV is between 0.0429 and 0.5850. Even though kNDVI has a smaller range compared with NIRV, the correlation between the corresponding KNDVI and canopy chlorophyll is still higher than the correlation between NIRV and canopy chlorophyll. This indicates that it is possible to find a range of sigma values to have a better correlation between kNDVI and most vegetation traits compared with the correlation between NIRV and corresponding vegetation traits. However, the detailed range of the sigma values is dependent on the property of vegetation traits as well as the distribution of observation.

### 3.2. Leaf traits over forests

#### 3.2.1. Study area and in-situ measurements

In this study, Harvard Forest (HF, 42.538N, 72.171W), located in the northeastern United States, is selected. This study site contains three tree species: red oak, red maple, and yellow birch. Annual precipitation and mean temperate are 1200 mm and 7.5 °C, respectively. In this area, leaf traits are being measured by using two leaves from red oak and red maples together with one leaf from yellow birch. Spectral properties are being measured via a spectro-radiometer ranging from 300 to 2500 nm where the spectral resolution is 3 nm for spectral till 700 nm and 7 nm in the remaining bands. The measured leaf traits parameters are chlorophyll concentration (including chlorophyll-a and chlorophyll-b, μ g/cm$^2$), carotenoids (μ g/cm$^2$), leaf mass per area (LMA, g/m$^2$), nitrogen concentration by mass (Nmass, %), and carbon concentration by mass (Cmass, %). Detailed information about the experiment setting can be found in Yang et al. (2016).

#### 3.2.2. Seasonal variability of leaf traits

Fig. 5 illustrates the time series of observation for carotenoids (*car*) and Cmass for red oak over 2002. It can be seen that the *car* increases from May 10th (3.36) on-wards and reached the peak value of 8.88 (μ g cm$^{-2}$) on August 7th, then starts to decrease till September 3rd. A Similar trend can be found for chlorophyll-a, chlorophyll-b, and total chlorophyll concentration for red oak in Fig. S5. In contrast, Nmass shows a reverse trend compared with other biophysical parameters. Similar to red oak, the red maple and yellow birch time series are also included in Fig. S6 and S7. It can be seen that yellow birch follows the same trend as red oak, which is indicated in Fig. S7. Nevertheless, a slight change can be indicated for red maple in Fig. S6, where chlorophyll-a, chlorophyll-b, total chlorophyll concentration, and carotenoids increase from May 1st onward and reach the peak value on June 26th, then starts to decrease. This contrasting behavior indicates that the phonology cycle of red maple is different from red oak and yellow birch.

#### 3.2.3. Applicability of kNDVI to estimate leaf traits

The identical configuration as in the previous experiment is used here. This experiment is carried out for three tree species individually, which results in 12 points for red oak and yellow birch as well as 13 points for red maple. The corresponding correlation coefficients between various leaf traits and different VIs for the three tree species are listed in Table 2. Similar to the previous section, scatter plots describing the relationship between biophysical parameters and VIs are also included (two examples are indicated in Fig. 6).

Table 2 indicates that for red oak and red maple, the kNDVI shows the most significant correlation for all the biophysical parameters followed by NIRv and NDVI, except NIRv, which shows a slightly higher

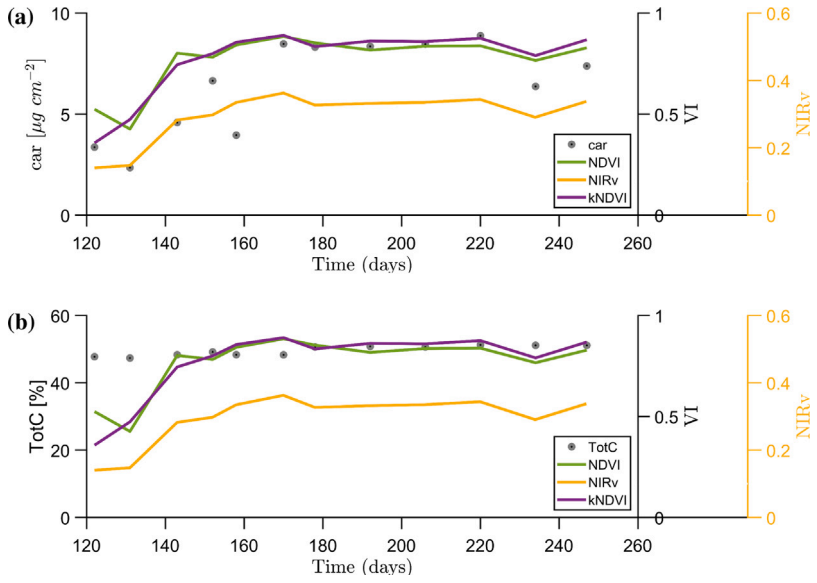


**Fig. 5.** Time series of carotenoids and Cmass traits for red oak.

**Table 2**
Correlation coefficient (absolute value) between various VIs and leaf traits for tree species ($p$-value is being indicated in the bracket).

| Tree species | Parameter | NDVI | NIRv | kNDVI |
|---|---|---|---|---|
| Red oak | Chl a | 0.789 (0.002) | 0.834 (0.001) | **0.844** (0.001) |
| | Chl b | 0.765 (0.004) | 0.805 (0.002) | **0.811** (0.001) |
| | Tot Chl | 0.784 (0.003) | 0.828 (0.001) | **0.837** (0.001) |
| | Car | 0.773 (0.003) | 0.809 (0.001) | **0.816** (0.001) |
| | LMA | 0.548 (0.065) | 0.661 (0.019) | **0.719** (0.008) |
| | TotC | 0.580 (0.048) | 0.619 (0.032) | **0.643** (0.024) |
| | TotN | 0.815 (0.001) | 0.861 (0) | **0.870** (0.001) |
| | median | 0.773 | 0.809 | **0.816** |
| Red maple | Chl a | 0.669 (0.012) | **0.750** (0.003) | 0.749 (0.003) |
| | Chl b | 0.652 (0.016) | 0.744 (0.004) | **0.747** (0.003) |
| | Tot Chl | 0.671 (0.012) | 0.756 (0.003) | **0.756** (0.003) |
| | Car | 0.417 (0.157) | 0.505 (0.078) | **0.520** (0.069) |
| | LMA | **0.770** (0.002) | 0.686 (0.010) | 0.637 (0.019) |
| | TotC | 0.262 (0.387) | 0.282 (0.350) | **0.327** (0.276) |
| | TotN | 0.782 (0.002) | 0.796 (0.001) | **0.834** (0) |
| | median | 0.669 | 0.744 | **0.747** |
| Yellow birch | Chl a | 0.745 (0.003) | **0.790** (0.001) | 0.766 (0.002) |
| | Chl b | 0.699 (0.008) | **0.754** (0.003) | 0.733 (0.004) |
| | Tot Chl | 0.736 (0.004) | **0.783** (0.002) | 0.760 (0.003) |
| | Car | 0.451 (0.122) | 0.567 (0.043) | **0.579** (0.038) |
| | LMA | 0.294 (0.329) | 0.655 (0.015) | **0.723** (0.005) |
| | TotC | 0.280 (0.353) | 0.424 (0.148) | **0.444** (0.129) |
| | TotN | 0.701 (0.008) | 0.874 (0) | **0.894** (0) |
| | median | 0.699 | **0.754** | 0.733 |

correlation with red maple compared with kNDVI. However, the difference (0.750 for NIRv and 0.749 for kNDVI) is negligible. For yellow birch, kNDVI shows the best correlation with carbon, nitrogen concentration, and carotenoids. Nonetheless, for biophysical parameters related to chlorophyll (Chlorophyll-a, Chlorophyll-b, and total Chlorophyll), NIRv shows the best correlation, followed by kNDVI and NDVI. NDVI does not show a high correlation with chlorophyll because its temporal variation is different from Chlorophyll as indicated in Fig. S7. The NDVI reaches peak value in late June, while Chlorophyll-a, Chlorophyll-b, and total Chlorophyll reach the maximum in early August. This result further confirms that the NDVI is not suitable for estimating the phenology cycle for the forest, which is in line with the findings of Zhang et al. (2022). It is also being observed in Fig. 6, which indicates NDVI has two clusters of values where the high values of NDVI around 0.8 are scattered while a quasi-linear relationship can

**Table 3**
Derivatives of vegetation indices for the calculation of the error propagation.

| | $\frac{\partial f}{\partial n}$ | $\frac{\partial f}{\partial r}$ |
|---|---|---|
| NDVI | $\frac{2r}{(n+r)^2}$ | $\frac{-2n}{(n+r)^2}$ |
| NIRv | $\frac{n^2+2nr-r^2}{(n+r)^2}$ | $\frac{-2n^2}{(n+r)^2}$ |
| kNDVI | $\frac{(n-r)}{2\sigma^2}\mathrm{sech}^2\left(\frac{(n-r)^2}{4\sigma^2}\right)$ | $-\frac{(n-r)}{2\sigma^2}\mathrm{sech}^2\left(\frac{(n-r)^2}{4\sigma^2}\right)$ |

be identified for the low values between 0.4 and 0.6. The correlation between kNDVI and Chlorophyll for yellow birch is slightly lower than the correlation between NIRv and Chlorophyll, as it is shown in Table 2. This is because the dynamic change of kNDVI (0.117–0.233) is slightly smaller than the change of NIRv (0.246–0.372) over the forest growth period, which can be seen in Fig. S7. This is different from the red oak case (Fig. S5), where the dynamic change of kNDVI (0.103–0.374) is much larger than the change of NIRv (0.140–0.363). This implies that for yellow birch, kNDVI may not perform better than NIRv in terms of the correlation with Chlorophyll a. The detailed reason for this should be further investigated, which might attribute to its biophysical or biochemical properties.

### 3.2.4. Error propagation for kNDVI with $\sigma$

Let us investigate the influence of $\sigma$ on error propagation by using the observed leaf reflectance for the three tree species. We compare the propagated error by the NDVI, NIRv, and kNDVI indices. Let us denote with $f(n,r)$ an arbitrary index acting on the NIR $n$ and red $r$ bands. Now, assuming independence between $n$ and $r$, we can calculate the propagated error with the variance formula:

$$s_f = \sqrt{\left(\frac{\partial f}{\partial n}\right)^2 s_n^2 + \left(\frac{\partial f}{\partial r}\right)^2 s_r^2} \tag{8}$$

where $s_f$ represents the standard deviation of the function index $f$, and $s_n$ ($s_r$) represents the standard deviation of $n$ ($r$). It is easy to compute the derivatives for each index according to Table 3.

In this experiment, we fixed the error $s_n = s_r = 0.004$ and calculated the indices and their propagated errors, see Fig. 7. We study the impact of different values of the $\sigma$ parameter (0.15, 0.22, and 0.3) on both the index and the propagated uncertainty.

It can be seen that the $\sigma$ nicely controls the sensitivity (step size) to vegetation density: the higher the $\sigma$, the lower the sensitivity (big steps

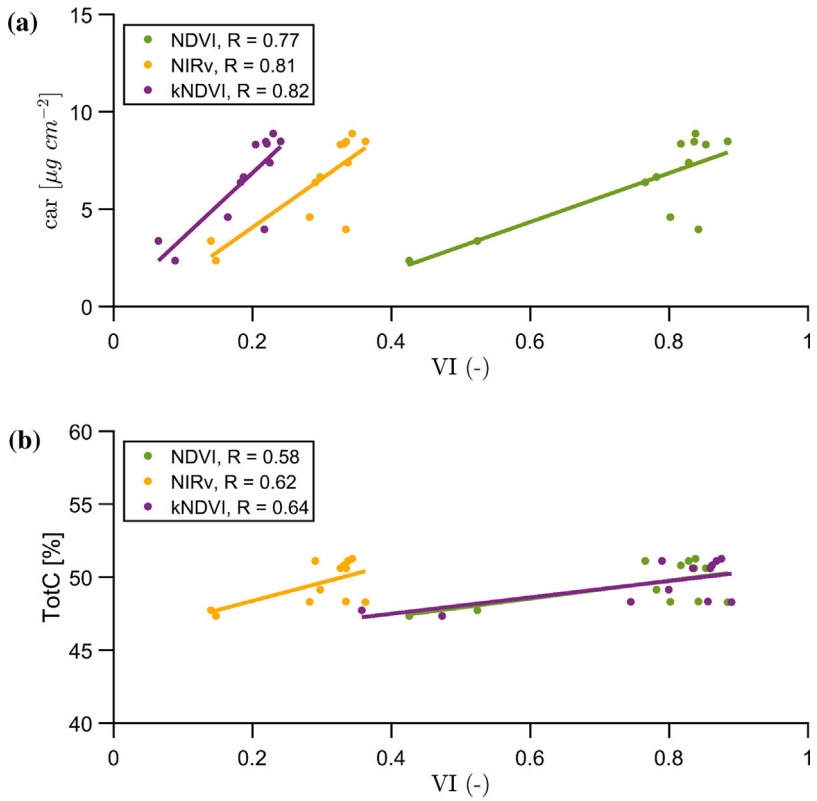


**Fig. 6.** Scatter plot for car (a) and totC (b) against VIs (NDVI, NIRv, kNDVI) for red oak.

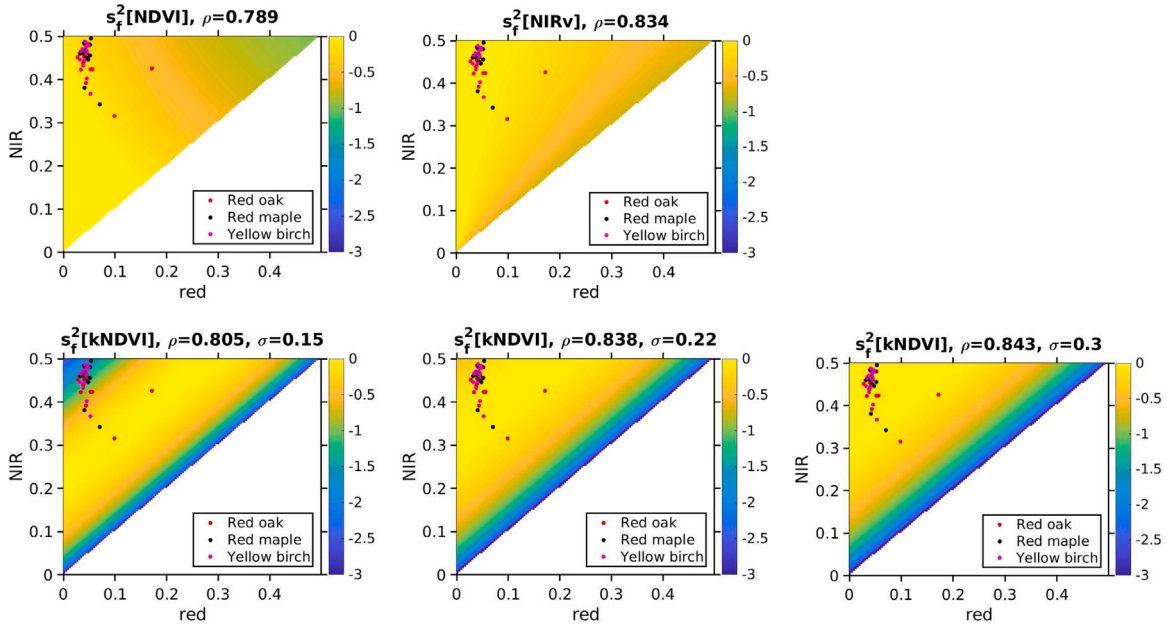


**Fig. 7.** Propagated error variances $s_f^2$ as a function of the NIR and red bands for the different vegetation indices. The input error was fixed to $s_n = s_r = 0.004$. The concrete *in situ* measurements for the three tree species (red oak, red maple, and yellow birch) are depicted for illustration purposes, and the corresponding correlation $\rho$ between the indices and measured Chl-a is shown.

are observed, shown in the last column of the lower panel in Fig. 7 with a broader greenness compared with the first and second column of the lower panel) which translates into lower kNDVI values.

Another interesting observation is that $\sigma$ clearly impacts the correlation $\rho$ of kNDVI with the vegetation trait at hand (here, we only show $\rho$ with the Chl-a parameter). Practitioners should pay attention and tune $\sigma$ for the specific applications. We hope these plots shed some light and intuition on how to tune $\sigma$: one should care about the land cover of interest and their spectral characteristics, how much sensitivity one aims to disentangle different levels (granularity) of species, and if one aims to avoid saturation, mixed-pixel or bias problems. One can also see that better correlations can be obtained with kNDVI compared to NDVI and NIRv, yet only with higher values of $\sigma$. Interestingly, the propagated uncertainty (note the log-scale in the right column of plots

in Fig. 7) by the kNDVI can be adjusted with properly tuning $\sigma$, which is a very interesting possibility to deal with different cover types (we here show that for forests, a low $\sigma = 0.15$ works best). In general, the region of lower propagated error (blueish, right) counteracts the high value of the index (yellowish, left), so a clear trade-off is observed.

### 3.3. Estimation of latent heat at flux tower level

#### 3.3.1. Dataset

FLUXNET 2015 dataset (Pastorello et al., 2020), which is a database that combines the data collected from multiple regional flux networks distributed over the globe and is being used in this study. It encompasses variables representing fluxes, meteorological, environmental, and soil status at half-hourly or hourly resolutions and provides the biggest compilation of in-situ observations to measure the exchanges of carbon dioxide, water vapor, and energy between the biosphere and atmosphere (Tramontana et al., 2016). For our analysis, a subset of 148 naturally vegetated areas is being selected according to the criteria of (1) no more than 50% of missing remotely sensed data due to cloud or snow contamination; (2) the temporal span of measurement should be more than four months to cover the whole phonological cycle.

#### 3.3.2. Applicability of kNDVI to estimate the latent heat flux

We evaluated the applicability to use kNDVI as a latent heat flux (LE) proxy using flux tower LE from eddy correlation from the FLUXNET 2015 dataset and compared its performance with NDVI, and NIRv. The correlation between LE and various VIs for different biomes are shown in Fig. 8. It can be observed that the kNDVI provides correlations with LE similar to (or better than the) other vegetation indices over all considered biomes and across all the flux tower sites. The per biome type assessment reveals that NDVI, NIRv and kNDVI show moderate to high correlation with LE and can be used as a proxy for LE for most of the biomes except evergreen broadleaf forest (EBF). For EBF, none of the considered VI performs as expected since a stronger saturation effect occurs in such an ecosystem Among the biomes, vegetation indices show the highest correlation with LE for mixed forests, with median correlations of 0.83,0.87 and 0.90 for NDVI, NIRv, and kNDVI, respectively. Overall, kNDVI generally outperforms the rest of the VIs to predict LE estimates over most of the considered biomes types. Only in the vegetation type savannas (SAV), the kNDVI performs slightly worse than NIRv. The kNDVI confirms its adaptive nature for each biome type but also globally shows an apparent gain (Fig. 9).

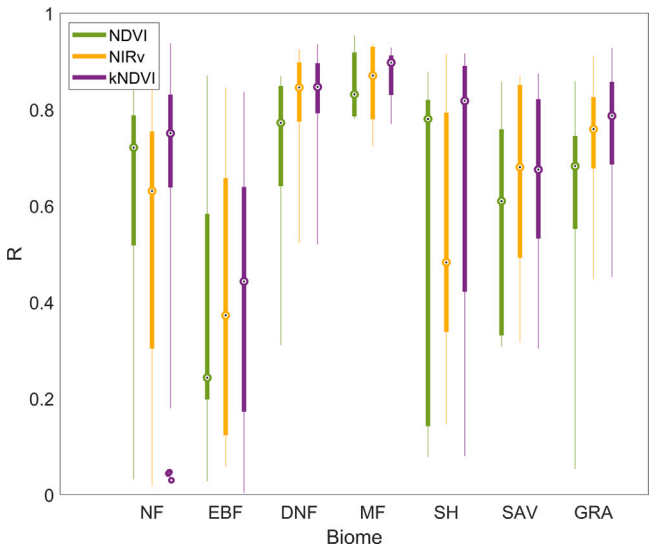


**Fig. 8.** Boxplots of correlations for all considered flux tower sites between LE and NDVI, NIRv, and kNDVI per biome type.

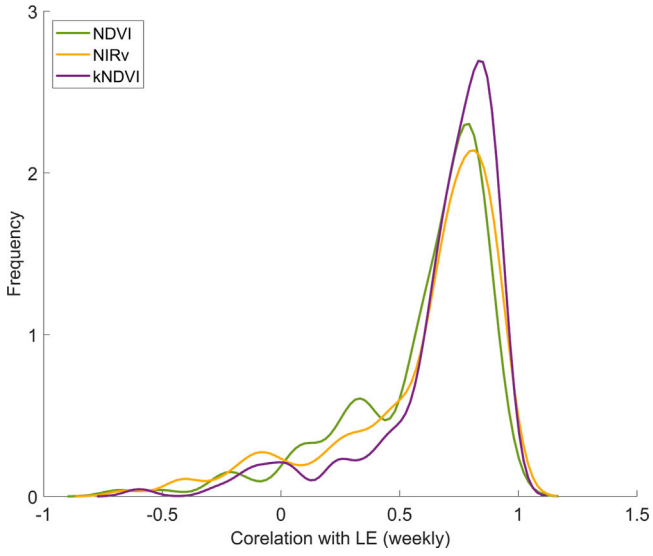


**Fig. 9.** Histogram of the correlation coefficient between the VIs and the parameters for the LE correlation computed over 169 FLUXNET sites.

## 4. Conclusions

We discussed the performance of the kNDVI (Camps-Valls et al., 2021), which generalizes the celebrated NDVI and other indices like NIRv. Generalization means that one has statistical guarantees of performance and that, in particular, one in principle may find a particular lengthscale ($\sigma$) of the RBF kernel function that retrieves the exact solution of NDVI (and NIRv). Therefore equal or better performance than these methods is expected. We clarified some aspects and properties of the kernel indices in general and kNDVI in particular by giving some applicable, practical prescriptions for proper usage. For example, we discussed the properties of generalization, boundedness, robustness to noise, and low error propagation. We also gave additional evidence of performance on a broader set of applications: estimating in-situ vegetation parameters (LAI, GPP, leaf, and canopy chlorophyll content, green and total LAI, and fAPAR) and latent heat at flux tower level. We confirmed the generally good performance of the index and highlighted its convenience in monitoring terrestrial ecosystems. To foster wider adoption of the new family of the index, we provide source code, data, and demos. The use of kernel methods to generalize NDVI could be extended to other indices, which will be the subject of further research.

## 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.

## Acknowledgments

We would like to acknowledge the Center of Land Management and Information Technologies and Carbon Sequestration Program at the University of Nebraska-Lincoln for providing the in-situ measurements. This work was partially supported by the European Research Council (ERC) under the ERC Synergy Grant USMILE (grant agreement 855187).

## Appendix A. Supplementary data

Supplementary material related to this article can be found online at https://doi.org/10.1016/j.isprsjprs.2022.12.019.

## References


Badgley, G., Field, C.B., Berry, J.A., 2017. Canopy near-infrared reflectance and terrestrial photosynthesis. Sci. Adv. 3 (3), e1602244.

Berni, J.A.J., Zarco-Tejada, P.J., Suárez, L., Fereres, E., 2009. Thermal and narrowband multispectral remote sensing for vegetation monitoring from an unmanned aerial vehicle. IEEE Trans. Geosci. Remote Sens. 47 (3), 722–738.

Camps-Valls, G., Bruzzone, L. (Eds.), 2009. Kernel Methods for Remote Sensing Data Analysis. Wiley & Sons, UK.

Camps-Valls, G., Campos-Taberner, M., Moreno-Martínez, Á., Walther, S., Duveiller, G., Cescatti, A., Mahecha, M.D., Muñoz-Marí, J., García-Haro, F.J., Guanter, L., et al., 2021. A unified vegetation index for quantifying the terrestrial biosphere. Sci. Adv. 7 (9), eabc7447.

Camps-Valls, G., Gómez-Chova, L., Muñoz-Marí, J., Vila-Francés, J., Calpe-Maravilla, J., 2006. Composite kernels for hyperspectral image classification. IEEE Geosci. Remote Sens. Lett. 3 (1), 93–97.

Daughtry, C.S.T., Walthall, C.L., Kim, M.S., De Colstoun, E.B., McMurtrey Iii, J.E., 2000. Estimating corn leaf chlorophyll concentration from leaf and canopy reflectance. Remote Sens. Environ. 74 (2), 229–239.

Gao, X., Huete, A.R., Ni, W., Miura, T., 2000. Optical–biophysical relationships of vegetation spectra without background contamination. Remote Sens. Environ. 74 (3), 609–620.

Gensheimer, J., Turner, A.J., Köhler, P., Frankenberg, C., Chen, J., 2021. A convolutional neural network for spatial downscaling of satellite-based solar-induced chlorophyll fluorescence (SIFnet). Biogeosciences Discuss. 1–25.

Gitelson, A.A., Gritz, Y., Merzlyak, M.N., 2003. Relationships between leaf chlorophyll content and spectral reflectance and algorithms for non-destructive chlorophyll assessment in higher plant leaves. J. Plant Physiol. 160 (3), 271–282.

Gorelick, N., Hancher, M., Dixon, M., Ilyushchenko, S., Thau, D., Moore, R., 2017. Google Earth engine: Planetary-scale geospatial analysis for everyone. Remote Sens. Environ. 202, 18–27.

Haboudane, D., Miller, J.R., Pattey, E., Zarco-Tejada, P.J., Strachan, I.B., 2004. Hyperspectral vegetation indices and novel algorithms for predicting green LAI of crop canopies: Modeling and validation in the context of precision agriculture. Remote Sens. Environ. 90 (3), 337–352.

Hase, N., Doktor, D., Rebmann, C., Dechant, B., Mollenhauer, H., Cuntz, M., 2022. Identifying the main drivers of the seasonal decline of near-infrared reflectance of a temperate deciduous forest. Agricult. Forest Meteorol. 313, 108746.

Huete, A.R., Liu, H., van Leeuwen, W.J.D., 1997. The use of vegetation indices in forested regions: Issues of linearity and saturation. In: IGARSS'97. 1997 IEEE International Geoscience and Remote Sensing Symposium Proceedings. Remote Sensing-a Scientific Vision for Sustainable Development, vol. 4. IEEE, pp. 1966–1968.

Kauth, R.J., Thomas, G.S., 1976. The tasselled cap–a graphic description of the spectral-temporal development of agricultural crops as seen by Landsat. In: LARS Symposia. p. 159.

Lancashire, P.D., Bleiholder, H., Boom, T.v.d., Langelüddeke, P., Stauss, R., Weber, E., Witzenberger, A., 1991. A uniform decimal code for growth stages of crops and weeds. Ann. Appl. Biol. 119 (3), 561–601.

Le Maire, G., Francois, C., Soudani, K., Berveiller, D., Pontailler, J.Y., Breda, N., Genet, H., Davi, H., Dufrane, E., 2008. Calibration and validation of hyperspectral indices for the estimation of broadleaved forest leaf chlorophyll content, leaf mass per area, leaf area index and leaf canopy biomass. Remote Sens. Environ. 112 (10), 3846–3864.

Myneni, R.B., Hall, F.G., Sellers, P.J., Marshak, A.L., 1995. Interpretation of spectral vegetation indexes. IEEE Trans. Geosci. Remote Sens. 33 (2), 481–486.

Pabon-Moreno, D.E., Migliavacca, M., Reichstein, M., Mahecha, M.D., 2022. On the potential of Sentinel-2 for estimating gross primary production. IEEE Trans. Geosci. Remote Sens..

Paruelo, J.M., Lauenroth, W.K., 1998. Interannual variability of NDVI and its relationship to climate for North American shrublands and grasslands. J. Biogeogr. 25 (4), 721–733.

Pastorello, G., Trotta, C., Canfora, E., Chu, H., Christianson, D., Cheah, Y.-W., Poindexter, C., Chen, J., Elbashandy, A., Humphrey, M., et al., 2020. The FLUXNET2015 dataset and the ONEFlux processing pipeline for eddy covariance data. Sci. Data 7 (1), 1–27.

Pei, Y., Dong, J., Zhang, Y., Yuan, W., Doughty, R., Yang, J., Zhou, D., Zhang, L., Xiao, X., 2022. Evolution of light use efficiency models: Improvement, uncertainties, and implications. Agricult. Forest Meteorol. 317, 108905.

Peng, Y., Gitelson, A.A., Keydan, G., Rundquist, D.C., WesleyMoses, 2011. Remote estimation of gross primary production in maize and support for a new paradigm based on total crop chlorophyll content. Remote Sens. Environ. 115 (4), 978–989.

Rojo-Álvarez, J., Martínez-Ramón, M., Muñoz Marí, J., Camps-Valls, G. (Eds.), 2018. Digital Signal Processing with Kernel Methods. Wiley & Sons, UK.

Rouse Jr., J., Haas, R.H., Schell, J.A., Deering, D.W., 1974. Monitoring vegetation systems in the great plains with ERTS.

Schölkopf, B., Smola, A., 2002. Learning with Kernels – Support Vector Machines, Regularization, Optimization and beyond. MIT Press Series.

Sellers, P.J., 1985. Canopy reflectance, photosynthesis and transpiration. Int. J. Remote Sens. 6 (8), 1335–1372.

Thenkabail, P.S., Smith, R.B., De Pauw, E., 2000. Hyperspectral vegetation indices and their relationships with agricultural crop characteristics. Remote Sens. Environ. 71 (2), 158–182.

Tramontana, G., Jung, M., Schwalm, C.R., Ichii, K., Camps-Valls, G., Ráduly, B., Reichstein, M., Arain, M.A., Cescatti, A., Kiely, G., et al., 2016. Predicting carbon dioxide and energy fluxes across global FLUXNET sites with regression algorithms. Biogeosciences 13 (14), 4291–4313.

Tucker, C.J., 1979. Red and photographic infrared linear combinations for monitoring vegetation. Remote Sens. Environ. 8 (2), 127–150.

Wang, X., Biederman, J.A., Knowles, J.F., Scott, R.L., Turner, A.J., Dannenberg, M.P., Köhler, P., Frankenberg, C., Litvak, M.E., Flerchinger, G.N., et al., 2022. Satellite solar-induced chlorophyll fluorescence and near-infrared reflectance capture complementary aspects of dryland vegetation productivity dynamics. Remote Sens. Environ. 270, 112858.

Wang, S., Zhang, L., Huang, C., Qiao, N., 2017. An NDVI-based vegetation phenology is improved to be more consistent with photosynthesis dynamics through applying a light use efficiency model over boreal high-latitude forests. Remote Sens. 9, 695.

Wang, S., Zhang, Y., Ju, W., Chen, J.M., Cescatti, A., Sardans, J., Janssens, I.A., Wu, M., Berry, J.A., Campbell, J.E., et al., 2021. Response to comments on "recent global decline of $CO_2$ fertilization effects on vegetation photosynthesis". Science 373 (6562), eabg7484.

Wu, W., 2014. The generalized difference vegetation index (GDVI) for dryland characterization. Remote Sens. 6 (2), 1211–1233.

Xue, J., Su, B., 2017. Significant remote sensing vegetation indices: A review of developments and applications. J. Sens. 2017.

Yang, X., Tang, J., Mustard, J.F., Wu, J., Zhao, K., Serbin, S., Lee, J.-E., 2016. Seasonal variability of multiple leaf traits captured by leaf spectroscopy at two temperate deciduous forests. Remote Sens. Environ. 179, 1–12.

Zeng, Y., Hao, D., Huete, A., Dechant, B., Berry, J., Chen, J.M., Joiner, J., Frankenberg, C., Bond-Lamberty, B., Ryu, Y., et al., 2022. Optical vegetation indices for monitoring terrestrial ecosystems globally. Nat. Rev. Earth Environ. 1–17.

Zhang, J., Xiao, J., Tong, X., Zhang, J., Meng, P., Li, J., Liu, P., Yu, P., 2022. NIRv and SIF better estimate phenology than NDVI and EVI: Effects of spring and autumn phenology on ecosystem production of planted forests. Agricult. Forest Meteorol. 315, 108819.

Zhu, L., Meng, J., Zhu, L., 2020. Applying geodetector to disentangle the contributions of natural and anthropogenic factors to NDVI variations in the middle reaches of the heihe river basin. Ecol. Indic. 117, 106545.