Sensitivity analysis of Sentinel-2 data for urban tree characterization using DART model

Théo Le Saint, Sidonie Lefebvre, Laurence Hubert-Moy, Jean Nabucet, Karine Adeline

Sensitivity analysis of Sentinel-2 data for urban tree characterization using DART model

Théo Le Saint*[a], Sidonie Lefevbre[b], Laurence Hubert-Moy [a], Jean Nabucet[a], Karine Adeline[c]

[a]UMR 6554 CNRS, LETG, University of Rennes, Place du Recteur Henri Le Moal, F-35000 Rennes, France; [b]ONERA-Department of Theoretical and Applied Optics, Paris-Saclay University, F-91123 Palaiseau, France; [c]ONERA-Department of Theoretical and Applied Optics, University of Toulouse, F-31055 Toulouse, France

ABSTRACT

The estimation of vegetation traits, which is essential to characterize the health of trees from remote sensing data, presents several challenges in urban environments, due to the topography of 3D buildings and associated shading, the spectral diversity of materials, or the variety of urban morphology. Moreover, the difficulty to estimate the vegetation traits increases with the decrease of spatial resolution, mixed pixels including information on trees and their environment. The objective of this study is to estimate the influence of tree-endogenous (chlorophyll, LAI...) and tree-exogenous (urban form, tree distance to buildings, street orientation, solar angles, materials types...) factors on the reflectance of Sentinel-2 pixels (10/20 m resolution). For this, a sensitivity analysis was carried out with the DART 3D radiative transfer model. First, a design of experiments was built using 15 variables describing the trees and their environment. Four urban 3D scenes that were elaborated based on the Local Climate Zone (LCZ) typology. For each of these urban 3D scenes, 3000 simulations were generated. Then, Sobol indices were computed to estimate the influence of each factor on the Sentinel-2 reflectance, more specifically on the 10 spectral bands and 8 vegetation indices correlated to vegetation traits. These experiments were conducted on isolated and aligned trees. In addition, the influence of the geo-registration uncertainty of the Sentinel-2 products was assessed in comparing the results obtained using a single tree-centered pixel with those using pixels offset from the tree. Results showed that Sentinel-2 data at 10 m resolution, NDVI et ARVI indices are the most relevant for the estimation of vegetation traits both for isolated and aligned trees, especially in LCZ 5 and 8, and in using a single tree-centered pixel approach.

Keywords: Urban tree, Vegetation traits, Sentinel-2, urban 3D scene, LCZ, DART, sensitivity analysis.

1. INTRODUCTION

In a context of climate change, urban tree is a central issue in planning policies [1] since it provides many vital ecosystem services [2], among them, temperature regulation through shading and evapotranspiration [3], carbon storage [4], and biodiversity preservation [5]. Thus, many benefits for human health and the living environment are conditioned by the presence of vegetation[6]. However, the urban environment is a context that can be unfavorable to the proper development of trees because of a large number of stress factors [7] that can originates from the underlying background having mechanical (soil compaction, artificialization increase with surrounding impervious materials) and chemical (soil contamination and lack of nutrients) disturbances, from the topography depending on the distance with manmade infrastructures (permanent or little light occlusion, constrained growth expansion), from the air (temperature, gas/aerosol pollution) or from the night lighting pollution. From the point of view of uses, trees in the city are therefore ambivalent: on the one hand, their presence is desirable and the benefits they are granted are numerous, on the other hand, the environment presents several constraints. Several studies show that these conditions have significant consequences for urban trees: the average lifespan is shorter for urban trees than for trees in rural areas [8], isolated trees and alignment trees suffer more stress than park trees [9], and the mortality rate of young trees is higher in cities [10].

These issues point to the need for large-scale monitoring of the dynamics and characteristics of urban vegetation. Through the literature, two indicators stand out as being particularly suitable for characterizing vegetation. Firstly, the chlorophyll content can be mentioned. Chlorophyll plays an essential role in photosynthesis and biomass production. In

addition, it is more sensitive to changes in external conditions than other pigments (e.g. carotenoids), so the chlorophyll content of leaves can be a good indicator of environmental stress, changes in temperature and humidity, as well as levels of pollutants in the air and soil [11]. Another indicator is the Leaf Area Index (LAI), which can identify the phenological stages of the tree, its photosynthetic potential. Chlorophyll content can be measured by taking leaf samples and extracting the pigments in the laboratory [12], or directly in situ with dedicated equipment (Dualex - FORCE A - Orsay - France, SPAD - Konica Minolta - Tokyo - Japan)[13]. On the other hand, the 3 main methods of indirect estimation of LAI are based on: hemispheric photography, LAI-2200 Plant Canopy Analyzer (LiCor - Lincoln - USA), TLS (terrestrial lidar scanning)[14]. For chlorophyll content as for LAI, these methods can be costly and time-consuming, and the implementation of large-scale monitoring tools (spatial and temporal) based on these methods seems difficult to implement.

A large number of studies are able to estimate these parameters by remote sensing, notably with Sentinel-2 images. However, these studies concern particular natural and landscape contexts, for example: tropical forests [15], temperate forests [16], mangroves [17] or even crops [18]. There are therefore few studies that apply these methods to the urban environment. The study of urban vegetation using remote sensing can take several forms, including inventory, assessment, biomass estimation, change detection, species classification and characterization [19]. These studies are based on data that, depending on the objectives, differ in terms of spatial, spectral and temporal resolution [20], [21]. The methods and data used vary even more according to the spatial scale of the object of study. If one wishes to characterize trees at the scale of individuals rather than at the scale of an urban park or forest, the data sources are restricted and the characterization of trees is essentially based on hyperspectral and LiDAR data [22]. In addition to spatial and spectral resolution, temporal resolution is one of the determining factors for the characterisation of vegetation and its dynamics[23]. Among the multi-spectral and hyperspectral satellites available in the operational phase (WorldView, RapidEye, Pleiades, QuickBird, enMAP, PRISMA, PlanetScope, etc.), Sentinel-2 is the only one that meets the following criteria: free images, satisfactory temporal resolution (weekly revisit). However, the spatial resolution (10m RGB/NIR and 20m RE/SWIR) of Sentinel-2 may be a limitation for the study of trees at the individual scale, the spectral resolution is also restricted. Moreover, the urban environment presents particularities that may constitute difficulties for the processing and analysis of satellite images. The strong geometric and volumetric heterogeneity due to the presence of buildings, and the associated shadowing, constitute a first constraint [24]. Urban spaces also present a significant spectral diversity of materials [25]. These cumulative factors have a strong influence on electromagnetic interactions and the signal perceived by a sensor. At 10 meters resolution, in an urban environment, the probability of obtaining a pure pixel in an image is very low. For a given tree, the reflectance values obtained in an image are influenced firstly by the tree itself: its structure, geometry, phenological stage, and foliage characteristics. Secondly, the urban context in which the tree is located, the distance of the tree from the buildings, the solar angles, the orientation of the street and the surrounding materials also have an influence on the pixel reflectance. The objective of this study is therefore to: (1) quantify the influence of exogenous tree parameters in several urban contexts, (2) establish a list of vegetation indices that are most influenced by the variables of interest (chlorophyll content, LAI) and finally (3) determine the most favorable context for monitoring these parameters. Radiative transfer models can be very powerful tools to better understand and analyze satellite images (ref). In this study, we use the DART radiative transfer model, developed at CESBIO (https://dart.omp.eu/#/). The DART model is a three-dimensional physical radiative transfer model that simulates the electromagnetic interaction between the Earth and the atmosphere in the visible to thermal infrared wavelengths. It allows the modeling of optical signals from urban and natural landscapes for any experimental configuration and instrumental specification. It is therefore possible to generate a scene that corresponds to a particular urban context, and to simulate the image acquired by a sensor whose resolutions (spatial and spectral) and acquisition angle are defined. From these images, a sensitivity analysis will be carried out to determine the influence of each input parameter on the generated images. The sensitivity analysis is based on the calculation Sobol indices (SI). The aim of this work is to evaluate the Sentinel-2 sensor's performance for characterizing urban trees. We are focusing on deciduous trees, as these are the species most commonly found in major European cities. We also focus on isolated trees and alignment trees. We also want to consider the different urban morphologies in the simulations, for which we are using the local climate zone framework [26]. This concept introduced in 2012 was originally dedicated to the study of urban climate and urban heat island, but it is increasingly used as a reference framework for studying urban vegetation [27]. The LCZs make it possible to take into account the heterogeneity of the types of building and the different proportions between built space, mineral surfaces and vegetal surfaces on a local scale. This makes it possible to have reference values to create 3D modeling models consistent with reality. The last item we want to investigate is more related to the specifications of the Sentinel-2 sensor and products. At 10 and 20 meters of resolution, when one wishes to extract pixel values for a tree (from a point or a polygon), the probability of obtaining a pure pixel is very low. In addition, Sentinel-2

products suffer from geolocation gaps (up to 12 meters between 2016 and 2021, and between 5 and 8 meters maximum from 2021 [28]). These elements constitute a source of bias and uncertainty. To simulate this variability, we have introduced an extraction window principle which can be centered on a point, or shifted with respect to the point according to a given angle. These elements therefore allow us to evaluate the sensitivity of Sentinel-2, according to two implementation scenarios (isolated tree or alignment tree), in several urban contexts (with LCZs), and to explore the variability of the measurements which can be carried out at the tree-scale (according to the different extraction windows).

2. MATERIALS AND METHODS

2.1 Dart parametrization and design of experiment

We consider here 2 different modeling scenarios. A first scenario with an isolated tree named **I** (SCI) and a second scenario with a tree alignment named **A** (SCA). These two scenarios are integrated into the 3D urban **models** corresponding to the 4 selected LCZS (2, 5, 6 and 8). We have studied the distribution of LCZs for European cities with more than 100,000 inhabitants based on the European map of LCZs [cite] and the administrative limits of European cities provided by EUROSTAT. The 5 most abundant types of LCZ are LCZ6 (42% of the area), LCZ9 (24%), LCZ8 (18.5%), LCZ5 (8.5%) and LCZ2 (2. 5%). To mark the difference between the LCZs, we have excluded the LCZ9 which is close to the LCZ6. The LCZ parameters were used to generate 3D models based on the following parameters: aspect ratio, fraction of built surface, fraction of impermeable surface, fraction of permeable surface and building height. The exact values are available in table 1.

Table 1. Information on video and audio files that can accompany a manuscript submission.

LCZ properties	LCZ 2	LCZ 5	LCZ 6	LCZ 8
Aspect ratio	1.375	0.525	0.525	0.2
Building fraction surface [%]	55	30	30	40
Impervious surface fraction [%]	40	40	35	45
Pervious surface fraction [%]	5	30	35	15
Height of roughtness element [m]	18	18	6.5	6.5

For SCI, the simulations were separated into 4 groups according to the 4 LCZs as presented in Figure 3. For each of the groups, 3000 simulations were performed. The fixes and variables parameters of these simulations are outlined in table 2. For SCA, in the same way as for SCI, the simulations were divided into 4 groups according to the LCZs. However, SCA represents an alignment of 5 trees with two different profiles (example for LCZ5 in figure 4). These two profiles are configured from the tree-endogenous parameters described in table2. The two profiles will therefore evolve with the simulations, but the distribution of the profiles within the alignment remains identical. The trees at the end of the alignment and the central tree are set with profile A, while the other trees are set with profile B. This configuration allows to integrate the heterogeneity of the trees within the alignments.

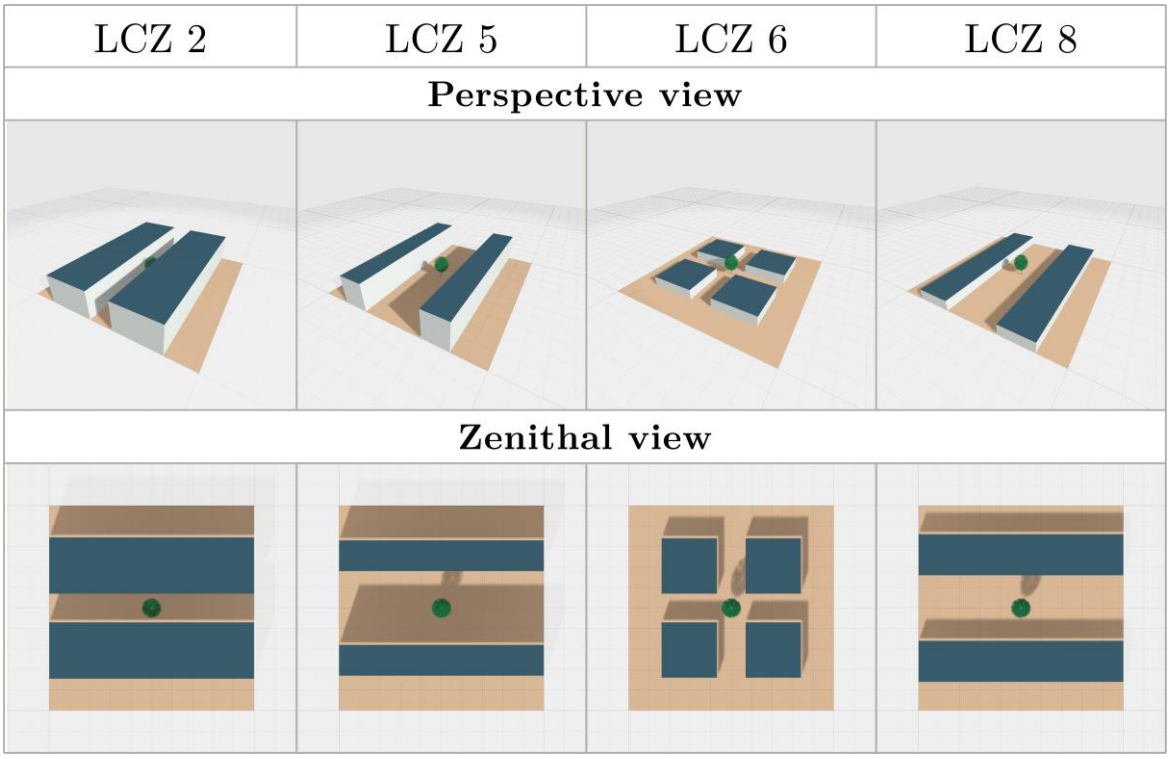

Figure 1. 3D mock-up for each LCZ in SCI.

As mentioned in section 2, the DART model has a large number of parameters to be set for the illumination, atmosphere, scene and sensor conditions. In the previous section, the main geometrical configuration for the scene has been justified and now we focus on describing the rest of the parameterization with parameters either fixed or variable (see Table 2). Two distinctions can be made beforehand:

- The first lies in the variability of the input parameters: indeed, some parameters have been fixed in the same way for all the simulations, while others vary according to predefined limits and distributions. Only the latter will be considered in the sensitivity analysis described in section 3.4.
- The second is semantic. The variable parameters have been grouped into two categories: the tree endogenous parameters (which have a direct influence on the tree canopy reflectance) and tree exogenous parameters (which can have an indirect influence on the tree canopy reflectance).

Table 2 lists the main DART parameters used in the simulations. They are described according to the section of the DART GUI to which they refer, their name, the category (exogenous or endogenous), their type (fixed or variable as mentioned above), the value of the parameter (a single value when it is a fixed parameter, a range of variation when it is a variable parameter). One of the columns mentions whether the parameters have been configured in the same way for SCI and SCA. Indeed, some parameters that were variable for SCI were then fixed for SCA because their influence was negligible.

Table 2. Dart parameterization: the section of the DART GUI (first column), name of the parameter (second column), parameter category (third column), parameter type, either fixed (F) or variable (V) (fourth column), parameter value (if F) or range of variation (if V) (fifth column), SCI and/or SCA configuration (sixth column) and comments (seventh column). Fixed parameters are used for SCI and SCA.

DART Section	Name	Category	Type	Values and range	SCI and / or SCA	Comment
Dart global settings	Light propagation mode	Exogenus	F	Bi-directional (DART-Lux)		
Sensor settings	Spectral bands	Exogenus	F	According to Sentinel-2 sensor		
	Zenith and azimuth angle	Exogenus	F	Zenith: 2.8 ° Azimuth: 182°		

	Spatial resolution	Exogenus	F	1m		Pixels are aggregated in post-processing
Direction input parameter	Hour	Exogenus	F	11h07 UTC		
	Day	Exogenus	F	Day 15		
	Month	Exogenus	V	From March to November	Both	
Atmosphere	Atmosphere model	Exogenus	F	Usstd76		
	Aerosol properties	Exogenus	F	Type : urban Optical depth : 1		
Scene optical properties	Roof	Exogenus	V	See figure 2	Both	
	Wall	Exogenus	V	See figure 2	Both	
	Impervious ground	Exogenus	V	See figure 2	Both	
	Pervious ground	Exogenus	V	See figure 2	Both	
Earth scene	Dimensions	Exogenus	F	100m x 100m		
	Latitude	Exogenus	F	48.1°		
	Longitude	Exogenus	F	-1.68°		
Tree implantation	Distance to building	Exogenus	V	LCZ2 and 6 : 5 – 6.5 [m] LCZ5 and 8 : 6-16 [m]	Both	
	Tree exposure	Exogenus	V	Shadow side or sunny side	Both	
	Street orientation	Exogenus	V	0, 45, 90, 135 [°]	Both	Anticlockwise, with 0° mean the street is oriented west-east
Tree	Geometric parameters	Endogenus	F	See table X		
	Leaf angle distribution	Endogenus	V	Plagiophile and planophile	Both	
	Leaf area index (*lai*)	Endogenus	V	0.1 and 2.5 [m2/m3]	Both	Defined via the leaf area density
	Clumping factor	Endogenus	V	0 - 50 %	SCI	
Leaf	Structure coefficient (*N*)	Endogenus	V	1.1 - 2.3 [arbitrary unit]	Both	
	Leaf chlorophylle content (*cab*)	Endogenus	V	5 - 60 [µg/cm2]	Both	Sum of chlorophyll a and chlorophyll b masses
	Carotenoid content (*car*)	Endogenus	V	2.5 - 25 [µg/cm2]	Both	
	Brown pigment	Endogenus	F	0 [arbitrary unit]		
	Anthocyane	Endogenus	F	0 [µg/cm2]		
	Equivalent water thickness	Endogenus	V	0.004 - 0.024 [cm]	SCI	
	Dry matter content	Endogenus	V	0.002 - 0.014 [g/cm2]	Both	

One of the particularities of the urban environment is the diversity of materials. The vast majority of these materials are artificial and of mineral origin. In this study, we have divided the materials into 4 classes: roof, wall, impermeable soil, permeable soil. Each of these classes will be assigned a list of possible reflectance spectra. Several spectral libraries of urban materials exist in the literature, with different levels of detail [29]. The use of very precise spectral libraries allows for a thorough analysis of the radiation balance, heat exchange, etc. In our case, we do not necessarily seek to be exhaustive, but rather to reproduce plausible contexts, and to cover a wide range of reflectance spectra that are well differentiated from one another. We use the following spectra available in the DART spectral library (this spectral library is available in the DART/database folder):

- Pervious ground: grass rye and grass dry

- Impervious ground: sandy stone, granit grey, asphalt grey and asphalt dark
- Wall: granit grey, concrete, concrete weathered, cement ochre, cement grey and brick
- Roof: aluminium roof, ceramic tile, zinc, gravel roof, cement tile and slate.

2.1 Simulated Sentinel-2 database

The DART outputs have been configured to match the Sentinel-2 configuration. Only the 10-meter (VIS and PIR) and 20-meter (RE, SWIR) bands were simulated. Central wavelengths and bandwidths were defined according to the characteristics of the MSI sensor. The angle of view was set to the average Sentinel-2 angle of view: 3° (with zenith = 0°). The spatial resolution has been set to 1 meter; this makes it possible to determine the grid for calculating pixel values in post-processing. Spatial aggregation to 10 (for visible and near-infrared bands) and 20 (for RE and SWIR bands) meters is done during post-processing. The DART outputs form a 100 x 100 pixels image. The post-processing consists of 3 steps: extraction of pixels of interest, aggregation at Sentinel-2 spatial resolutions and calculation of vegetation indices.
To extract pixel values, we consider 4 extraction windows of 10x10 meters:

- one window centered on the tree (window 1)
- 3 windows shifted by 5 meters with respect to the center of the tree: in the direction of the alignment (at 0°) (window 2), at 45° (window 3) and at 90° (window 4)

The extraction of the pixel values of the spectral bands at 20 meters is performed from a 20x20m. However, the 20x20m window is not centered on the 10x10m window (because this does not correspond to a real configuration between the two Sentinel-2 resolutions). They are arranged so that the 10x10m window corresponds to the lower left quarter of the 20x20m window. The pixel values of the DART output are then extracted for the window area and averaged. In this way, for each DART image, we can extract 1 or more spectra according to the different extraction windows. 14 vegetation indices (VI) have been calculated and associated with the LUTs. The list of indices is available in the appendix.

Table 3. List of computed vegetation indices (VIs), columns present the type (simple, normalized or multibands), the name, the different bands used (MS for multispectral and VIS for visible), the resolution of the indices and the references.

Type	Nom	Bands	Resolution	References
Simple	Simple ratio	MS	10m	[33]
Normalized	Green Leaf Index	VIS	10m	[34]
	Normalized green–blue difference index	VIS	10m	
	Normalized green–red difference index	VIS	10m	[35]
	Red Green Blue Vegetation Index	VIS	10m	[36]
	Atmosphéric resistant vegetation index	MS	10m	[37]
	Normalized Vegetation Index	MS	10m	[38]
	Red edge NDVI	MS	20m	[39]
	Sentinel-2 LAI Index	MS	20m	[40]
Multibands	Sentinel-2 Triangular Vegetation Index	MS	20m	[41]
	Modified Chlorophyll Absosrption in Reflectance Index 2	MS	10m	[42]
	Transformed Chlorophyll Absorption Reflectance Index	MS	20m	[43]
	Optimized Soil Adjusted Vegetation Index	MS	10m	[44]

Finally, the spectra and the VIs are associated with the input parameters of the DART simulations to create a look-up table (LUT)

For SCI and SCA, 24 LUTs are generated:

- 4 for SCI: one for each LCZ with the window n°1
- 16 for SCA depending on the LCZs and the extraction window

2.2 Sensitivity analysis

Sensitivity analysis is used to evaluate the sensitivity of a model output regarding the variation of input parameters. In our case, the model outputs are the spectral bands and the vegetation indices and the input parameters are the variable parameters described in table 2. The 'Sobol' method is a global sensitivity analysis based on variance decomposition. For our study, the Sobol indices (SI) have several advantages over other types of sensitivity indices such as HSIC, Fourrier or Morris indices. First, they measure the interaction effect between input parameters, which is important when variables are interdependent (for example: it is the case for the 3 parameters: orientation, distance to the building and exposure, which will all 3 have an influence on the proportion of shadow in the pixel). Then, they are based on a decomposition of the overall variance, which makes it possible to separate the contribution of each parameter, measure the relative importance of each and rank the contributions. This is interesting here in order to rank the vegetation indices according to the contribution associated with the targeted parameters such as *cab* or *lai*. Finally, they can be computed efficiently for a large number of input parameters, which is important in complex applications where many parameters need to be taken into account. In our case, we have several LUTs, let's take the example of the LUT corresponding to SCI for the LCZ2. This LUT is a dataset of 3000 rows and 34 columns. Each record corresponds to a single simulation. The variables include: the input parameters of the DART simulations (16 parameters), the spectral values (10 spectral bands) and the vegetation index values (8 indices). Sobol indices are used to study the sensitivity of a mathematical model [30]. In our case, the simulations are based on a complex physical model (DART), so it is necessary to calculate a metamodel. This metamodel is then used as a reference model for the calculation of Sobol indices. We used a gaussian process regression (GPR) to compute several multivariate metamodels. For each output parameter: spectral bands and vegetation indices, a metamodel was computed from the input parameters. All metamodels have been validated with a test dataset representing 25% of the data. For each metamodel, the R squared was calculated, ranging from 0.84 to 0.99. Then the sobol indices (SI) were performed using the Saltelli algorithm [31] avalaible in the OpenTurns python library. Metamodel and Sobol indices calculation were realized with the OpenTurns python library.

3 RESULTS

3.1 Sensitivity indices for isolated tree scenario

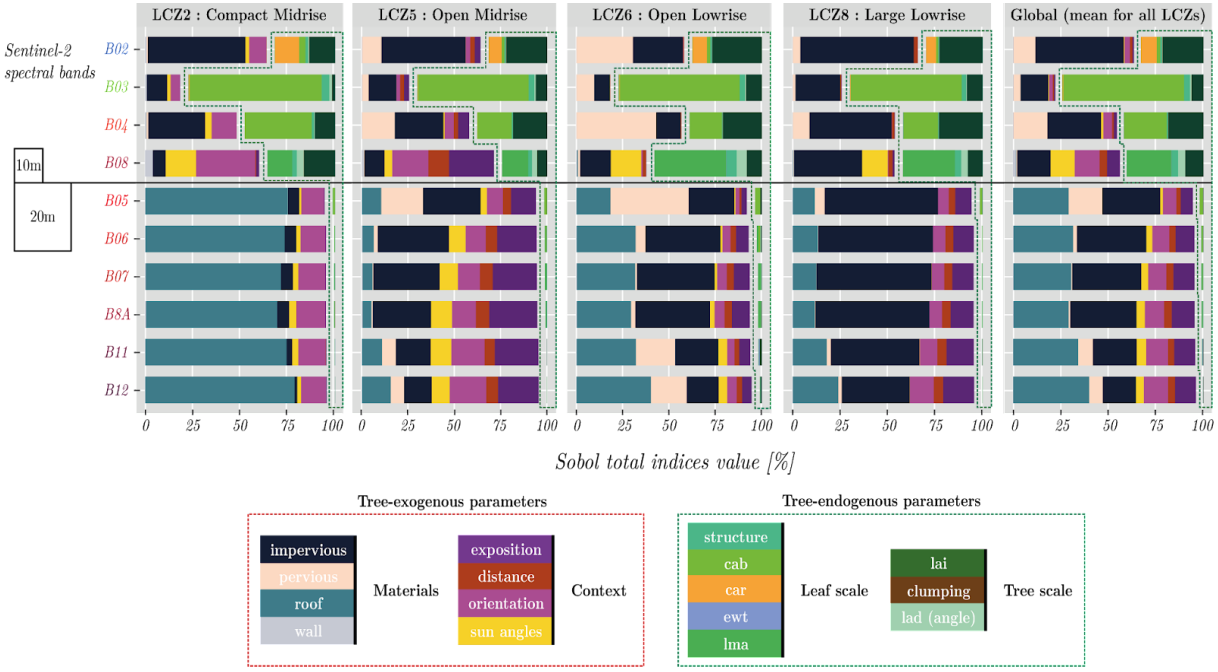

Figure 2. Graphical representation of Sobol Index (SI) values for each spectral band and LCZ. Each line corresponds to a spectral band, and each colored band corresponds to the contribution value of the parameter. The columns correspond to the different LCZs.

This part of the results concerns the simulations carried out for the 4 LCZs with the SCI as presented in Figure 1. Firstly, the parameters can be grouped into 4 categories: tree scale, leaf scale, contextual parameters and finally material types. The first two categories are endogenous to the tree while the last two are exogenous. A first element is common to all LCZs: the prominent influence of the exogenous parameters on the spectral bands at 20 meters, as can be seen from the SI values for the global case (average of the 4 LCZs), the part of the tree-endogenous parameters is negligible with SI values below 1%. However, the distribution between the different contributions evolve according to the LCZ. For LCZ 2 the major contribution comes from the roof materials while for LCZ5, the influence of the optical properties of the soil is more important. The contribution of the contextual parameters (all of which have an influence on the shadow cast on the pixel) is also important. These elements show that, in an urban canyon, the width of the street therefore has an influence on the signal, however, the small proportion of vegetation (one tree) in a 20x20m pixel has no influence on the signal. In general, the distribution of the different factors is related to the fraction of impervious surface: when it is important (75% for LCZ8), the influence of impervious materials dominates, while when this fraction is more restricted (55% for LCZ5 and 50% for LCZ6), the contribution of pervious materials increases. This first element limits the use of Sentinel-2 images, although the bands at 20 m have a useful spectral richness for the characterization of vegetation, in urban areas, they are too influenced by exogenous parameters.

Consider now the bands at 10 m. In general, we can identify the same phenomenon as for the 20m bands: the influence of soil materials and shading remains very important and the contribution varies according to the LCZ. Band B2 is, to a small extent, the only one to be influenced by carotene content. *car* may be an interesting indicator, especially during the scenic period, however, the blue band is predominantly influenced by soil materials. The B03 band is the most influenced by the *cab* with 73%, 62%, 68% and 61% for LCZ2, 5, 6, and 8 respectively. Contributions to band B08 are more heterogeneous with strong sensitivity to street orientation, exposure and solar angles for LCZ2 and 5. For LCZ6 and 8, the lma parameter has a non-negligible influence (with respectively 40% and 28%). The contribution of the ewt parameter (related to leaf water content) is totally negligible in this context and at these resolutions. Its influence on the signal is almost null, even in the SWIR bands with the water absorption peak.

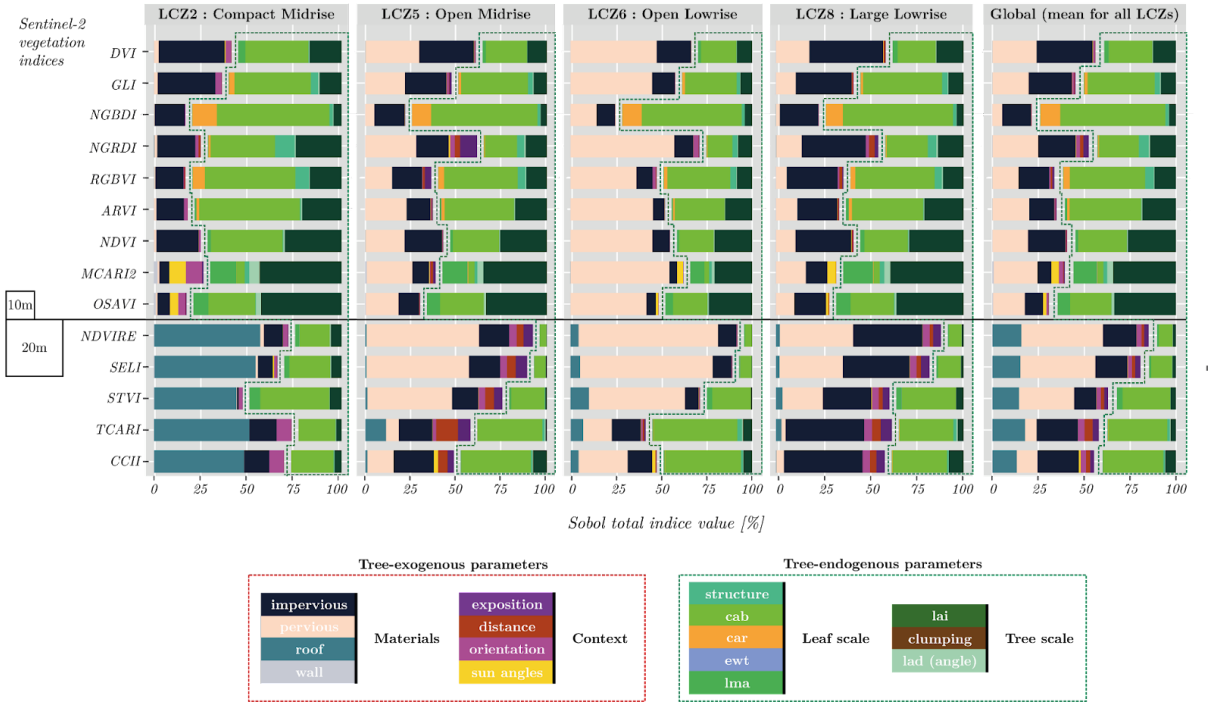

Figure 3. Graphical representation of Sobol Index (SI) values for each VIs and LCZ. Each line corresponds to a VI, and each colored band corresponds to the contribution value of the parameter. The columns correspond to the different LCZs.

If we first look at the global values (average for the 4 LCZs), a large number of indices are mostly influenced by tree-endogenous parameters. For the indices calculated from the 10m bands, only the DVI is mostly influenced by exogenous parameters. Concerning the indices calculated at 20m, we observe the same phenomenon as for the spectral bands: the contribution of exogenous parameters is higher, and in a significant proportion for NDVIRE and SELI.
The influence of the LCZ is also characterized in the same way as for the spectral bands: LCZ2 is marked by the importance of the roof materials, LCZ5 is the one most influenced by contextual factors (distance to the building, orientation, exposure). The SI values for impervious and pervious materials reflect the proportion of impervious to pervious surface inherent to each LCZ. However, with vegetation indices, the contribution of permeable soil can be very important since it is also vegetation, this is particularly the case for LCZ6.
When considering a tree with a crown of 10 meters in diameter, its surface area when viewed from above represents 78m². Compared to a pixel of 20m, the proportion of the pixel occupied by the tree is 19.5%. Just to facilitate the reading of the results, we can make the simplistic assumption that there is a linear relationship between the proportion of trees in a pixel (in terms of area) and the SI values of the tree-endogenous parameters. In this respect, we can mention the performance of TCARI, especially for LCZ6 where the proportion of grass is more important, but where the SI values for the endogenous parameters reach 60%.
The point made previously that the 20m bands seemed unusable is thus qualified here by the performance of the TCARI, especially when compared to the performance of other indices, notably those at 10m, for LCZ6.
If we now look at the indices calculated from the bands at 10 meters, we can read the results according to 3 categories: the normalized indices calculated from the visible bands, the normalized multispectral indices and the complex multispectral indices.

- Normalized indices with visible bands (GLI, NGBDI, NGRDI, RGBDI):

Being based on the visible bands, these 4 indices are more oriented towards the pigment content of the leaves, especially thanks to the absorption peaks of the *cab* in the blue and red bands. At first, we can note that these 4 indices are influenced by the endogenous factors but in a different way, if we take the global values for the endogenous factors (average of the 4 LCZs), we obtain the following classification: NDGBI (77.5%), RGBVI (64%), GLI (52%) and NGRVI (45%). If we focus on the best: NDGBI is the index most influenced by *cab*, and this for all LCZs. NDGBI is also the most constant index.

- Normalized indices with multispectral infrareds bands (ARVI, NDVI):

The results for ARVI and NDVI are quite similar. On the global SI (last column of figure 9), we can note the major contribution of *cab* (40% for ARVI and 27% for NDVI), a sensitivity to *lai* is also introduced with the use of B08 band (20% for ARVI and 27% for NDVI). These two indices are also very sensitive to the proportion of pervious soil, which can be identified by comparing the SI values between LCZ2 and 6. For example, between these two LCZs, SI values for *cab* decrease from 55% to 28% for ARVI and from 40% to 20% for NDVI.

- Complexes multi spectral infrareds indices (MCARI2, OSAVI) :

These two indices show the greatest heterogeneity in the contribution of endogenous parameters to the tree. Initially, the MCARI formula was developed to respond to *cab* variation, but Daughtry et al. showed that it was also sensitive to *lai* and *lai* - *cab* interaction. MCARI2 was introduced to maintain maximum sensitivity to *lai* while reducing the influence of *cab* variation. This pattern is found in this study: MCARI2 has the lowest SI value for *cab* with 3.5% overall, and the highest sensitivity to *lai* with 35% overall. The contribution of the LMA is also not negligible with 13% overall, it is the only index (with the OSAVI in a lesser proportion) to be sensitive to this parameter. However, as before, we find the sensitivity to pervious soil.

3.2 Sensitivity indices for alignment tree scenario

This second part of the results presents the sensitivity analysis for the SCA scenario. Figure 6s shows the SI values for the vegetation indices, for each combination of LCZ and extraction window (with, for each LCZs, the average of the SI values for the 4 windows). Vegetation index results obtained on 20-meter bands have been excluded from the results, as they show too low a contribution from tree-endogenous parameters (see section 3.1). If we first consider extraction window n°1 (centered on the tree). The SCI and SCA results are very similar. The same influential parameters are found for the different VIs, and the effect of LCZs is identical. When a pixel is centered on a tree, there is therefore no prominent influence of the type of layout, whether the tree is isolated or within an alignment. Nevertheless, the probability of obtaining a tree-centered pixel in a Sentinel-2 image is very low.

Most of the pixels extracted from Sentinel-2 images correspond to the different extraction windows presented in the method. Firstly, for window 2 (which corresponds to a window straddling two trees, the percentage of total tree

occupancy is 78%, as for window 1), the parameters of both tree profiles are represented in the contributions, and in equivalent proportions. For the NGBDI in LCZ5, for example, the SI values for the parameters *cab* of tree A and *cab* of tree B are 28% and 30% respectively. For all the VIs, there is an equivalent division between the contributions of tree A and tree B. However, for most VIs, the total contribution of tree-endogenous parameters decreases between window 1 and window 2. This effect is all the more significant in LCZs 5 and 6, which have more permeable ground surfaces, and for most VIs, this parameter makes the greatest contribution. This result shows that even with an identical area of tree vegetation in the pixel, the contribution of this vegetation to index values can change, depending on soil type and siting context.

Windows 3 and 4 are offset from the alignment. As a result, for most VIs, the contribution of tree-exogenous parameters increases considerably. For each LCZ, the contributions of material types are the most important, with those of the roof for LCZ2, mostly those of permeable soil for LCZ5 and 6, and those of impermeable soil for LCZ8. Contextual parameters also make a notable contribution. Mainly in LCZ 2, 5 and 8. Indeed, as soon as the pixel is no longer centered on the tree, the influence on the signal of the shadow cast by the buildings or by the tree itself is more important. These elements are consistent with the properties of each LCZ (see Table 1 and Figure 1 in section 2.1).

The Sobol indices can be used to identify the various parameters influencing the VIs. In addition, we can also determine which indices are most influenced by the parameters of interest in each LCZ. Concerning LCZ2, if we look at the overall SI values (average of all windows), NGBDI and ARVI obtain the best contributions from the *cab* parameter (*cab* of tree A + *cab* of tree B) with 52% and 45% respectively. For *lai*, MCARI2 and OSAVI are the most influenced with 32% and 27% respectively. Similar patterns are found for the other LCZs, with NGBDI and ARVI for the *cab* parameter and MCARI2 and OSAVI for *lai*. However, these contributions may be relatively low in certain contexts, notably for windows 3 and 4 for LCZ6. For ARVI, for example, the SI value for the cab parameter, window 4 and LCZ6, is only 6%. This trend confirms the influence of underlying vegetation on most VIs. Finally, the NGBDI does not follow the same trend as the other indices and is less influenced by the permeable soil parameter than the other indices, in the case of mixed pixels (windows 3 and 4). Indeed, the contribution of the *cab* parameter to the NGBDI in LCZ6, window 4, reaches 40%.

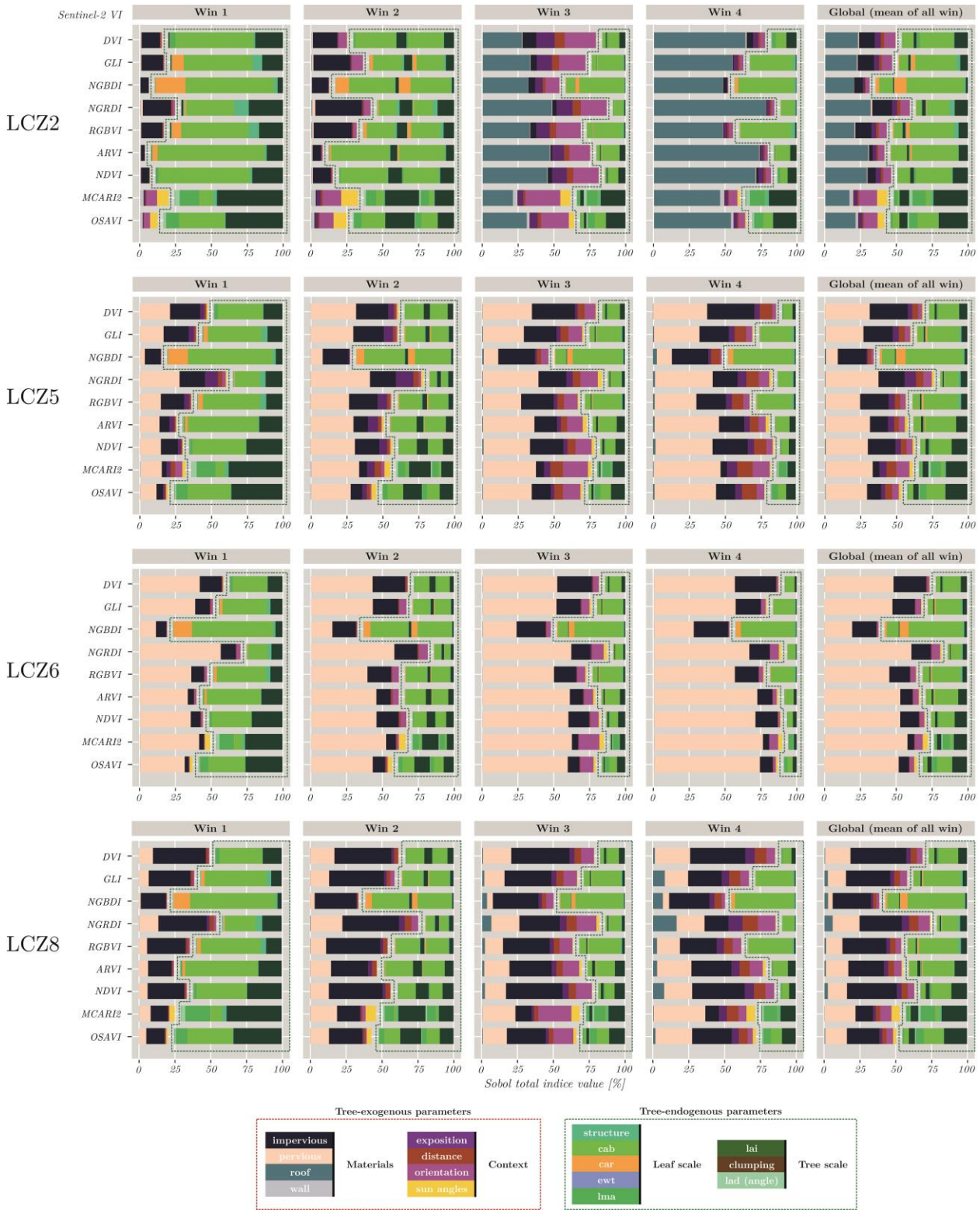

Figure 4. Graphical representation of Sobol Index (SI) values for each spectral band, LCZ and extraction window. Each line corresponds to a VI, and each colored band corresponds to the contribution value of the parameter. The columns correspond to the different windows

4 DISCUSSION AND CONCLUSION

This work highlights the capabilities and limitations of the Sentinel-2 sensor for characterizing trees in urban environments. Simulations are based on simplified representations of trees planted in an urban context. A number of factors have been introduced into the simulations to cover and summarize a wide range of cases. The LCZ nomenclature enables us to reconstruct different urban morphologies in terms of volumes, layout, building geometry and the proportions of different types of material. On the other hand, the DART model is based on a large number of parameters and shows good results in urban environments [32]. Firstly, the sensor parameters are almost exhaustive, in terms of spatial and spectral resolution as well as acquisition conditions (viewing and solar angles). This allows us to simulate realistic Sentinel-2 images. Next, scene characteristics can be defined with great precision, from the scale of the atmosphere, the urban scene, the tree, then the leaf (based on the PROSPECT model). However, in order to be able to simulate Sentinel-2 images, certain choices had to be made when parameterizing the DART model. The type of atmosphere and aerosols chemical composition is fixed. Aerosols are known to be very present and variable in urban areas [33], and have a significant influence on reflectance, particularly in the blue wavelengths. Atmospheric parameterization could be improved in two ways: by integrating atmospheric type and aerosol optical depth into the sensitivity analysis as a variable parameter, or by integrating atmospheric composition from observations made in urban environments at different times of the year to cover different climatic conditions.

On another scale, the use of a single model for tree geometry may be questioned, but in this study, it allows us to generalize and consider an average tree, which will correspond to the majority of cases encountered. Although the management of urban tree heritages introduces types of cutting that are often regular, tending to standardize tree geometry and profiles. However, there is a diversity of tree types and species present in the city, and this diversity leads to a pluralism of tree heights, dimensions and foliage densities. 3D models derived from LIDAR measurements could be integrated into the simulations to obtain a more accurate 3D model of the tree.

The sensitivity analysis presented here provides a ranking of the various input parameters influencing the Sentinel-2 signal. This method is reproduced with the same input parameters in several urban contexts (LCZ) and according to several extraction windows. The extraction windows allow us to restore two aspects: the uncertainty linked to the georeferencing of Sentinel-2 products, and the probability (inherent in the size of the study object in relation to the image resolution) of obtaining a mixed pixel in an urban environment with Sentinel-2. We have shown in the results that the two main factors, LCZ and extraction window, have a considerable impact on the Sentinel-2 signal. More specifically, we have seen with the different LCZs that the underlying vegetation is very influential if we want to study trees, and the size of the buildings and the location of the tree in relation to them will determine the proportion of shadow cast in the pixel, and the presence of shadow in the pixel has a significant effect on reflectance values, particularly for the B08 band and the resulting VIs. On the other hand, the contributions of tree-endogenous parameters fall drastically as one moves away from the tree and the proportion of tree vegetation in the pixel is reduced (e.g. between window 1 and window 4). In this context, the use of Sentinel-2 20-meter bands seems unlikely for isolated or aligned trees (unless they have a very large diameter, greater than that modelled in our simulations). Among the best-performing indices, we can notably cite the NGBDI and the ARVI, which are the most influenced by the cab parameter. However, the high performance of the NGBDI index gives rise to reflection on the use of the visible bands (in particular the B02 band) in an urban environment because of their sensitivity to aerosols. This index represents the slope of the spectrum between the wavelengths of blue and green. According to the sensitivity analysis, the angle of this slope is strongly influenced by the presence of chlorophyll in the leaves. However, if vegetation pixels are considered, the absolute reflectance value of B02 and B03 have a low amplitude. Therefore, a small variation in the reflectance values of an image (the reflectance may vary in the presence of aerosols, or according to the atmospheric correction algorithm used) can induce a strong variation in the index value. For the lai parameter, MCARI2 and OSAVI give the best performance. By identifying the most influential parameters on the Sentinel-2 signal, the sensitivity analysis initially makes it possible to classify the best performing indices. But on the other hand, it also makes it possible to highlight the most influential exogenous parameters.

For each vegetation index, we can therefore list and quantify the sources of uncertainty. As part of a study, many of these parameters can be known, identified and spatialized through auxiliary data. The uncertainty related to these parameters can thus be integrated into machine learning models in order to improve their performance. Finally, this work presents a simplified and reproducible method for modelling tree vegetation in an urban environment. It can also be applied to other sensors, to assess their performance in characterizing urban trees and/or estimating biophysical parameters such as chlorophyll content or LAI. This type of study can help determine the spectral and spatial resolution best suited to these objectives.

REFERENCES

[1] F. Xu, J. Yan, S. Heremans, et B. Somers, « Pan-European urban green space dynamics: A view from space between 1990 and 2015 », *Landscape and Urban Planning*, vol. 226, p. 104477, oct. 2022, doi: 10.1016/j.landurbplan.2022.104477.

[2] Y. Andersson-Sköld *et al.*, « An integrated method for assessing climate-related risks and adaptation alternatives in urban areas », *Climate Risk Management*, vol. 7, p. 31-50, janv. 2015, doi: 10.1016/j.crm.2015.01.003.

[3] P. Bolund et S. Hunhammar, « Ecosystem services in urban areas », *Ecological Economics*, vol. 29, nº 2, p. 293-301, mai 1999, doi: 10.1016/S0921-8009(99)00013-0.

[4] D. J. Nowak et D. E. Crane, « Carbon storage and sequestration by urban trees in the USA », *Environmental Pollution*, vol. 116, nº 3, p. 381-389, mars 2002, doi: 10.1016/S0269-7491(01)00214-7.

[5] E. Andersson, S. Barthel, et K. Ahrné, « Measuring Social–Ecological Dynamics Behind the Generation of Ecosystem Services », *Ecological Applications*, vol. 17, nº 5, p. 1267-1278, 2007, doi: 10.1890/06-1116.1.

[6] K. L. Wolf, S. T. Lam, J. K. McKeen, G. R. A. Richardson, M. van den Bosch, et A. C. Bardekjian, « Urban Trees and Human Health: A Scoping Review », *Int J Environ Res Public Health*, vol. 17, nº 12, p. 4371, juin 2020, doi: 10.3390/ijerph17124371.

[7] M. Czaja, A. Kołton, et P. Muras, « The Complex Issue of Urban Trees—Stress Factor Accumulation and Ecological Service Possibilities », *Forests*, vol. 11, nº 9, Art. nº 9, sept. 2020, doi: 10.3390/f11090932.

[8] A. Sæbø *et al.*, « The Selection of Plant Materials for Street Trees, Park Trees and Urban Woodland », in *Urban Forests and Trees: A Reference Book*, C. Konijnendijk, K. Nilsson, T. Randrup, et J. Schipperijn, Éd., Berlin, Heidelberg: Springer, 2005, p. 257-280. doi: 10.1007/3-540-27684-X_11.

[9] B. Ma, R. J. Hauer, J. Östberg, A. K. Koeser, H. Wei, et C. Xu, « A global basis of urban tree inventories: What comes first the inventory or the program », *Urban Forestry & Urban Greening*, vol. 60, p. 127087, mai 2021, doi: 10.1016/j.ufug.2021.127087.

[10] D. Hilbert, L. Roman, A. Koeser, J. Vogt, et N. van Doorn, « Urban Tree Mortality: A Literature Review », *AUF*, vol. 45, nº 5, sept. 2019, doi: 10.48044/jauf.2019.015.

[11] F. Talebzadeh et C. Valeo, « Evaluating the Effects of Environmental Stress on Leaf Chlorophyll Content as an Index for Tree Health », *IOP Conf. Ser.: Earth Environ. Sci.*, vol. 1006, nº 1, p. 012007, avr. 2022, doi: 10.1088/1755-1315/1006/1/012007.

[12] H. K. LICHTENTHALER et A. R. WELLBURN, « Determinations of total carotenoids and chlorophylls a and b of leaf extracts in different solvents », *Biochemical Society Transactions*, vol. 11, nº 5, p. 591-592, oct. 1983, doi: 10.1042/bst0110591.

[13] R. Casa, F. Castaldi, S. Pascucci, et S. Pignatti, « Chlorophyll estimation in field crops: an assessment of handheld leaf meters and spectral reflectance measurements », *The Journal of Agricultural Science*, vol. 153, nº 5, p. 876-890, juill. 2015, doi: 10.1017/S0021859614000483.

[14] S. Wei *et al.*, « An assessment study of three indirect methods for estimating leaf area density and leaf area index of individual trees », *Agricultural and Forest Meteorology*, vol. 292-293, p. 108101, oct. 2020, doi: 10.1016/j.agrformet.2020.108101.

[15] H. Padalia, S. K. Sinha, V. Bhave, N. K. Trivedi, et A. Senthil Kumar, « Estimating canopy LAI and chlorophyll of tropical forest plantation (North India) using Sentinel-2 data », *Advances in Space Research*, vol. 65, nº 1, p. 458-469, janv. 2020, doi: 10.1016/j.asr.2019.09.023.

[16] Y. Li *et al.*, « Fine-scale leaf chlorophyll distribution across a deciduous forest through two-step model inversion from Sentinel-2 data », *Remote Sensing of Environment*, vol. 264, p. 112618, oct. 2021, doi: 10.1016/j.rse.2021.112618.

[17] J. Zhen *et al.*, « Mapping leaf chlorophyll content of mangrove forests with Sentinel-2 images of four periods », *International Journal of Applied Earth Observation and Geoinformation*, vol. 102, p. 102387, oct. 2021, doi: 10.1016/j.jag.2021.102387.

[18] J. G. P. W. Clevers, L. Kooistra, et M. M. M. Van den Brande, « Using Sentinel-2 Data for Retrieving LAI and Leaf and Canopy Chlorophyll Content of a Potato Crop », *Remote Sensing*, vol. 9, nº 5, Art. nº 5, mai 2017, doi: 10.3390/rs9050405.

[19] A. R. Shahtahmassebi *et al.*, « Remote sensing of urban green spaces: A review », *Urban Forestry & Urban Greening*, vol. 57, p. 126946, janv. 2021, doi: 10.1016/j.ufug.2020.126946.

[20] L. Velasquez-Camacho, A. Cardil, M. Mohan, M. Etxegarai, G. Anzaldi, et S. de-Miguel, « Remotely Sensed Tree Characterization in Urban Areas: A Review », *Remote Sensing*, vol. 13, nº 23, Art. nº 23, janv. 2021, doi: 10.3390/rs13234889.

[21] J. Degerickx, D. A. Roberts, J. P. McFadden, M. Hermy, et B. Somers, « Urban tree health assessment using airborne hyperspectral and LiDAR imagery », *International Journal of Applied Earth Observation and Geoinformation*, vol. 73, p. 26-38, déc. 2018, doi: 10.1016/j.jag.2018.05.021.

[22] R. Näsi *et al.*, « Remote sensing of bark beetle damage in urban forests at individual tree level using a novel hyperspectral camera from UAV and aircraft », *Urban Forestry & Urban Greening*, vol. 30, p. 72-83, mars 2018, doi: 10.1016/j.ufug.2018.01.010.

[23] C. Granero-Belinchon, K. Adeline, et X. Briottet, « Impact of the number of dates and their sampling on a NDVI time series reconstruction methodology to monitor urban trees with Venμs satellite », *International Journal of Applied Earth Observation and Geoinformation*, vol. 95, p. 102257, mars 2021, doi: 10.1016/j.jag.2020.102257.

[24] T. Zhou, H. Fu, C. Sun, et S. Wang, « Shadow Detection and Compensation from Remote Sensing Images under Complex Urban Conditions », *Remote Sensing*, vol. 13, nº 4, Art. nº 4, janv. 2021, doi: 10.3390/rs13040699.

[25] M. Jilge, U. Heiden, C. Neumann, et H. Feilhauer, « Gradients in urban material composition: A new concept to map cities with spaceborne imaging spectroscopy data », *Remote Sensing of Environment*, vol. 223, p. 179-193, mars 2019, doi: 10.1016/j.rse.2019.01.007.

[26] I. D. Stewart et T. R. Oke, « Local Climate Zones for Urban Temperature Studies », *Bulletin of the American Meteorological Society*, vol. 93, nº 12, p. 1879-1900, déc. 2012, doi: 10.1175/BAMS-D-11-00019.1.

[27] C. Zhao, Q. Weng, Y. Wang, Z. Hu, et C. Wu, « Use of local climate zones to assess the spatiotemporal variations of urban vegetation phenology in Austin, Texas, USA », *GIScience & Remote Sensing*, vol. 59, nº 1, p. 393-409, déc. 2022, doi: 10.1080/15481603.2022.2033485.

[28] « Forthcoming deployment of the Copernicus Sentinel-2 products geometric refinement », *Sentinel Online*. https://copernicus.eu/-/forthcoming-deployment-of-the-copernicus-sentinel-2-products-geometric-refinement (consulté le 13 août 2023).

[29] S. Kotthaus, T. E. L. Smith, M. J. Wooster, et C. S. B. Grimmond, « Derivation of an urban materials spectral library through emittance and reflectance spectroscopy », *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 94, p. 194-212, août 2014, doi: 10.1016/j.isprsjprs.2014.05.005.

[30] B. Iooss et P. Lemaître, « A Review on Global Sensitivity Analysis Methods », in *Uncertainty Management in Simulation-Optimization of Complex Systems: Algorithms and Applications*, G. Dellino et C. Meloni, Éd., in Operations Research/Computer Science Interfaces Series. Boston, MA: Springer US, 2015, p. 101-122. doi: 10.1007/978-1-4899-7547-8_5.

[31] A. Saltelli, Éd., *Sensitivity analysis in practice: a guide to assessing scientific models*. Hoboken, NJ: Wiley, 2004.

[32] Z. Zhen *et al.*, « DART: a 3D radiative transfer model for urban studies », in *2023 Joint Urban Remote Sensing Event (JURSE)*, mai 2023, p. 1-4. doi: 10.1109/JURSE57346.2023.10144212.

[33] Z.-C. Zeng *et al.*, « Remote sensing of angular scattering effect of aerosols in a North American megacity », *Remote Sensing of Environment*, vol. 242, p. 111760, juin 2020, doi: 10.1016/j.rse.2020.111760.