

Good Practice Guidelines for UAV-based Surface Reflectance Validation



Authors: Origo, N., Kalacska, M., Arroyo-Mora, J. P., Soffer, R., Smigaj, M., Brede, B., De Los Reyes, R., Koehler, C., Pflug, B., Niro, F., Sinclair, M., Ong, C., Lau, I., Ramsay, R., Byrne, G., Morris, H., Suomalainen, J., Kooistra, L., Brell, M., Randall, C., Latini, D., Wilk, J., Raqueno, N., Gerace, A., Bauch, T., Eon, R., Dash, J., Camacho, F., Sanchez-Zapero, J., Martinez, E., Brown, L., Morrone, R., Mota, B., Gillespie, J., Merrington, A., Van der Zee, J., Petracca, I., Thankappan, M., Ormane, R., Scholes, M., Honkavaara, E., Broomhall, M., González Piqueras, J. & Boccia, V.

SRIX4Veg2 participants: Ong, C., Lau, I., Sparrow, B., Lucieer, A., Origo, N., Maier, S., Whiteside, T., Barnetson, J., Broomhall, M., Haynes, R., Morris, H., Sinclair, M., Koh, J., Franz, T., Byrne, G., Franks, L., Lewis, M., Levick, T., Levick, S. & Johnson, S.

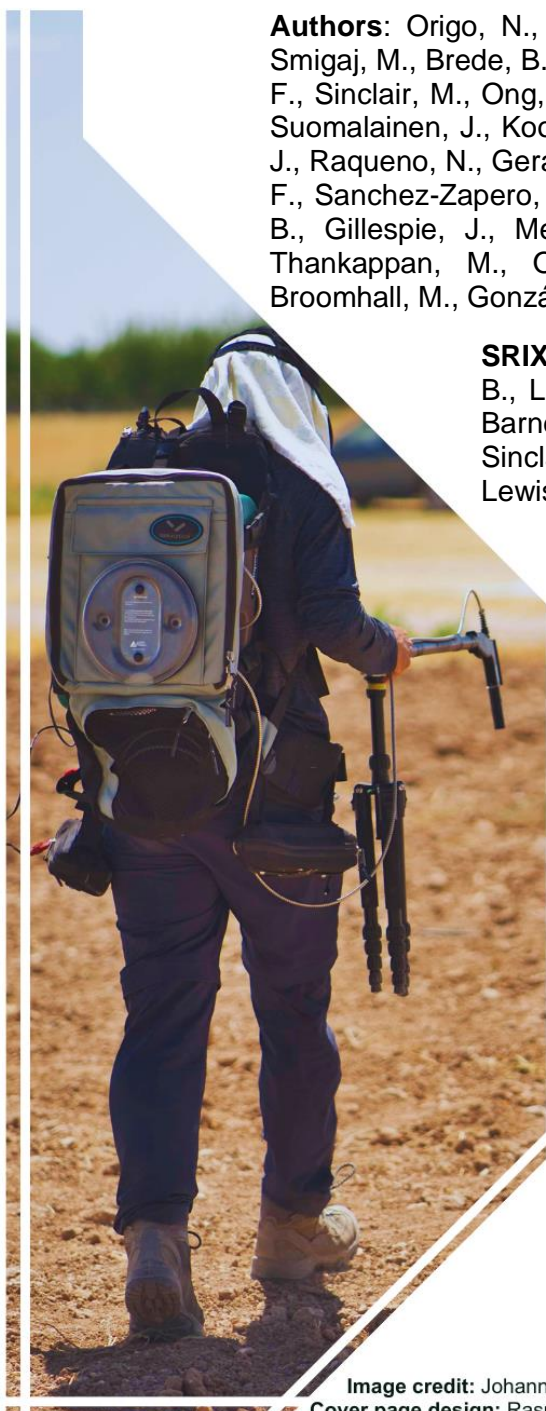


Image credit: Johannes Wilk
Cover page design: Rasma Ormane



Australian Government
Geoscience Australia



List of abbreviations

ACIX	Atmospheric Correction Intercomparison Exercise
AERONET	Aerosol Robotic Network
AFOV	Angular Field Of View
AGL	Above Ground Level
APU	Accuracy Precision and Uncertainty
ARD	Analysis Ready Data
ARD-VP	ARD validation protocol
BOA	Bottom-Of-Atmosphere
BCRF	Bi-Conical Reflectance Factor
BIPM	Bureau International des Poids et Mesures
BRDF	Bi-directional Reflectance Distribution Function
BVLOS	Beyond Visual Line Of Sight
CEOS	Committee on Earth Observation Satellites
CORS	Continuously Operating Reference Station
DEA	Digital Earth Australia
DGPS	Differential GPS
DLS	Downwelling Light Sensor
DSM	Digital Surface Model
ECV	Essential Climate Variable
ELM	Empirical Line Method
FOV	Field Of View
FRM	Fiducial Reference Measurement
FWHM	Full Width at Half Maximum
GA	Geoscience Australia
GCOS	Global Climate Observing System
GCP	Ground Control Point

GIS	Geographical Information System
GNSS	Global Navigation Satellite System
GSD	Ground Sampling Distance
GUM	Guide to the expression of Uncertainty in Measurement
HCRF	Hemispherical-Conical Reflectance Factor
HSI	Hyperspectral Imager
HYPERNETS	Hyperspectral Network
IMU	Inertial Measurement Unit
ISO	International Organisation for Standardization
JCGM	Joint Committee for Guides in Metrology
LPV	Land Product Validation
MC	Monte Carlo
MODIS	Moderate-resolution Imaging Spectrometer
MTOM	Maximum Take-Off Mass
MTOW	Maximum Take-Off Weight
NTRIP	Networked Transport of RTCM via Internet Protocol
PD	Pixel Duplication
PDE	Path Definition Error
PL	Pixel Loss
PPK	Post-Processing Kinematic
PSF	Point Spread Function
PSM	Panel Substitution Method
RTC	Real-Time Correction
RTCM	Radio Technical Commission for Maritime Services
RTK	Real-Time Kinematic
SAA	Solar Azimuth Angle
SfM	Structure from Motion

SLSTR	Sea and Land Surface Temperature Radiometer
SNR	Signal-to-Noise Ratio
SR	Surface Reflectance
SRIX4Veg	Surface Reflectance Intercomparison eXercise for Vegetation
SZA	Solar Zenith Angle
UAS	Unoccupied Aerial System
UAV	Unoccupied Aerial Vehicle
UTM	Universal Transverse Mercator
WGCV	Working Group on Calibration and Validation
WGS	World Geodetic System

Table of Contents

List of abbreviations	2
1 INTRODUCTION	7
1.1 Goal of this document	7
1.2 Importance of surface reflectance.....	7
1.3 The role of CEOS WGCV	8
1.4 Importance of and need for UAV-based surface reflectance validation	8
1.5 SR requirements.....	10
2 DEFINITIONS.....	12
3 GENERAL CONSIDERATION FOR SATELLITE SR PRODUCT VALIDATION	15
3.1 Quantities	15
3.2 Surface.....	15
3.3 Anisotropy	16
3.4 Illumination considerations	16
3.5 Spectral considerations	17
4 GENERAL STRATEGIES FOR COLLECTION OF SR REFERENCE DATASETS	18
4.1 Case study: CEOS-WGCV-endorsed surface reflectance protocol for ground-based measurements	21
5 GOOD PRACTICES FOR UAV-BASED SR REFERENCES (CONDENSED VERSION)	24
5.1 Sensor.....	24
5.2 UAS.....	24
5.3 Environmental conditions.....	24
5.4 Flight planning	25
5.5 Supplementary data.....	25
5.6 Post-processing	26
6 GOOD PRACTICES FOR UAV-BASED SR REFERENCES	27
6.1 Reflectance methods.....	27
6.2 Instrument types and supplementary UAV system components	28
6.3 Geometry.....	30
6.3.1 GNSS	31
6.4 Flight planning	32
6.4.1 UAV flight conditions	33

6.4.2	UAS flight performance	33
6.4.3	Flight planning software	36
6.4.4	Advanced applications – estimation of spatial coverage	39
6.5	Reference targets	43
6.5.1	Panel substitution method (PSM).....	43
6.5.2	Reference target characteristics.....	47
6.6	Post-processing	48
6.6.1	VZA filtering	49
6.6.2	Hyperspectral data aggregation	50
6.7	Calibration and characterisation	50
6.8	Uncertainty characterisation and propagation	53
6.8.1	Radiometry	58
6.8.2	Geometry	60
6.8.3	Image quality	60
6.8.4	Simulations	61
7	<i>STRATEGY FOR SR VALIDATION BASED ON UAV SR MEASUREMENTS</i>	62
8	<i>CONCLUSIONS.....</i>	64
9	<i>References.....</i>	65

1 INTRODUCTION

1.1 Goal of this document

This document aims to establish a good practices protocol for validating satellite surface reflectance products using Unoccupied Aerial Vehicle (UAV)-based spectroradiometers. The protocol originates from an intercomparison campaign conducted in July 2022 in Barrax (Spain) and March 2024 in Calperum (Australia) under the auspices of the Committee for Earth Observation Satellites (CEOS) – Working Group on Calibration and Validation (WGCV) endorsed surface reflectance intercomparison exercise for vegetation (SRIX4Veg/2) initiatives. Emphasis will be placed on adopting metrological practices and aligning with the CEOS Fiducial Reference Measurements (CEOS-FRM) concept (Goryl et al., 2023). Particular attention will be devoted to assessing the uncertainty budget associated with these measurements, ensuring their suitability for validating satellite-based surface reflectance products. The formulation of this protocol is crucial for CEOS WGCV as it addresses the need for best practices for UAV-based surface reflectance measurements. The protocol is aimed at the users of UAV-mounted instruments capable of validating surface reflectance products and those wishing to enter this field. The document is a community effort, and the authors have a wide variety of research experience which is mainly focused on land applications. This document is owned by CEOS-WGCV and will be updated periodically.

1.2 Importance of surface reflectance

Surface Reflectance (SR) is generally defined as the fraction of incoming solar radiation reflected from Earth's surface for specific observation and environmental conditions. Different reflectance quantities can be specified for each condition, as in Schaepman-Strub et al., (2006). If surface reflectance measurements are properly interpreted using rigorous terminology, they can provide critical information about terrestrial and aquatic surface properties and their temporal changes.

In the terrestrial domain, which is the focus of this document, the most critical parameter is the surface albedo, which is a key forcing parameter controlling the planetary radiative energy budget and the partitioning of radiative energy between the atmosphere and the surface. More broadly, surface reflectance serves as a crucial input for retrieving a wide range of terrestrial essential climate variables (ECVs) related to vegetation, land cover and land use, snow, and fire disturbances.

Surface reflectance is also a fundamental surface-level parameter derived from satellite optical imaging sensors and it is commonly provided to users as an intermediate Level 2 product. Harmonisation of surface reflectance products across missions is therefore a fundamental step in the pathway towards enhanced interoperability (Niro et al., 2021). In this regard, the CEOS Analysis Ready Data ([CEOS-ARD](#)) framework offers comprehensive guidelines on how to ensure

that surface reflectance from different sensors is derived and presented to users in a uniform manner, with standardised content and metadata, hence facilitating their integrated usage.

1.3 The role of CEOS WGCV

The role of the CEOS WGCV is to ensure long-term confidence in the accuracy and quality of satellite-based Earth Observation data and products and provide a forum for the exchange of information about calibration and validation and associated coordination, and cooperative activities.

CEOS WGCV is the first CEOS point of contact for the international user-community as far as calibration and validation, system technical information and Earth Observation (EO) quality processes are concerned. To this end, CEOS WGCV addresses the need to standardise ways of combining data from different sources to ensure the interoperability required for the effective use of existing and future EO systems.

To achieve these goals, CEOS WGCV is organised into various sub-groups, each dedicated to specific and focused areas of interest. These sub-groups act as discussion forums to address topics on a comprehensive scientific and technical basis, allowing for action on selected areas of interest.

The Land Product Validation ([LPV](#)) subgroup within CEOS WGCV is specifically focused on land products. Its mission is to coordinate the quantitative validation of satellite-derived terrestrial ECVs. The emphasis is on standardised intercomparison and validation across products from diverse satellites, algorithms, and agencies. The main outcomes of the LPV subgroup are a set of community-endorsed [protocols](#) providing, for each ECV, the measurement good practices for in-situ data collection, spatial sampling, upscaling approaches, and recommended methods for comparison with correlative satellite data, along with associated validation metrics.

1.4 Importance of and need for UAV-based surface reflectance validation

Within the LPV subgroup, there is a dedicated focus area for surface radiation variables. Previous efforts in this domain focused on albedo, resulting in a good practices [protocol](#) issued in 2019. This protocol covered the validation of satellite albedo products using existing ground-based networks, airborne campaigns, and satellite product intercomparison. More recently, an online validation tool with associated reference datasets and standardised reporting was developed for the validation of land surface albedo products (Sánchez-Zapero et al., 2023). As a result of this tool, and considering the availability of community good practices, globally representative validation datasets and a rigorous estimation of uncertainties in satellite data, the albedo product was recently upgraded to validation “Stage 4”, the highest maturity level in the [LPV validation hierarchy](#).

While the validation maturity of satellite albedo products is very high, little attention has been devoted so far to the validation of its precursor product, surface reflectance. There are several reasons for that. Firstly, the LPV subgroup has historically focused on terrestrial ECVs, as defined by the Global Climate Observing System (GCOS) (GCOS-244/272, 2022). Surface reflectance, strictly speaking, is not an ECV, since its generic form, the Hemispherical Directional Reflectance Factor (HDRF), which is commonly adopted in several satellite products, is not an intrinsic property of the surface, but it depends on the observation and environmental conditions (solar geometry and direct/diffuse-to-total irradiance influenced by cloud and atmospheric properties). Secondly, surface reflectance is an intermediate product in the satellite processing chain, spanning the raw sensor data (Level 0) to the calibrated top-of-atmosphere radiances/reflectances (Level 1), and extending to the derived surface bio-geophysical parameters (Level 2). Consequently, in the frame of CEOS-WGCV, surface reflectance holds a transitional role, situated between the thematic domain of the Infrared and Visible Optical Sensor (IVOS) subgroup, which focuses on Level 1 products, and the Land Product Validation (LPV) subgroup, which focuses on derived Level 2 products.

This lack of maturity of surface reflectance products has been identified as crucial in recent years, stimulating various agencies, under the auspices of CEOS-WGCV, to allocate efforts to enhance the validation maturity of the surface reflectance product. Notable endeavours in this regard include the Atmospheric Correction Intercomparison Exercise (ACIX) (Doxani et al., 2018), the HYPERNETS network (Goyens et al., 2021), and the Fiducial Reference Measurements for Vegetation (FRM4Veg) project (Origo et al., 2020), and in 2018 CEOS-WGCV endorsed a community agreed SR validation ground-based protocol (Malthus et al., 2019a) which was developed by the Australian Earth Observation community for validation of Landsat 7 and Sentinel-2 under the Digital Earth Australia (DEA) project (Byrne et al., 2021). As a result, DEA are now able to publish ongoing validation metrics as they apply to the Landsat and Sentinel-2 SR products. A summary of the WGCV endorsed protocols and results are given in Malthus et al., (2019b) (section 4).

ACIX predominantly addressed the challenge of surface reflectance validation through an indirect approach, relying on reference synthetic surface reflectances estimated across a representative set of Aerosol Robotic Network (AERONET) sites. The HYPERNETS network aimed to establish the first global network of automated spectroradiometers, providing a reference dataset for validating land and water satellite surface reflectance products. As part of the FRM4Veg project, the surface reflectance validation issue was approached from a metrological standpoint, involving a detailed analysis of the uncertainties associated with ground- and satellite-based measurements and their comparison. The SRIX4Veg initiative naturally extends the activities of FRM4Veg by specifically addressing UAV-based measurements. This protocol is therefore

essential to CEOS-WGCV for enhancing the maturity of validation practices for surface reflectance validation.

The ability to collect reference data for SR validation has notably expanded in the past years because of the increased availability and accessibility of drone-based multi- and hyper-spectral optical systems. UAV platforms offer numerous advantages for validation purposes. They ensure extensive and cost-effective geographical coverage, especially in challenging and hard-to-reach locations, unlike traditional operator-based field surveys. Additionally, UAVs enable denser spatial sampling, addressing the gap in the upscaling process from ground-based point measurements to satellite pixel-scale resolution.

It is important to recognise that UAVs come with their own set of challenges such as trade-offs in relation to the payload that can be carried and for how long, regulatory control (in many countries UAVs must be registered with the local aviation authority and flights may only be permitted in certain locations due to airspace rules, etc.; Stöcker et al., 2017), and that the cost associated with UAVs can be cheaper than traditional methods but there is a transition cost in terms of the cost of the UAV and payload, training, trial-and-error, and regular maintenance to ensure airworthiness. Despite the evident benefits of UAVs, there is currently no community agreed-upon protocol for UAV-based SR validation. Conversely, several good practices have been published for validating surface reflectance using traditional field spectroscopy measurements, as in Held et al., (2015). We consider this lack of consensus a significant drawback, and the primary objective of this document is to address this shortcoming.

1.5 SR requirements

GCOS does not have specific requirements for all surface reflectance quantities, however, explicit criteria are defined for the albedo product. In the latest GCOS documents, requirements for the albedo product were revised compared to previous versions (GCOS-245, 2022). Notably, more stringent criteria were established for the spatial resolution (target = 250 m, goal = 10 m) and measurement uncertainties (target = 5%, goal = 3%), reflecting advances in the resolution and accuracy of the most recent data assimilation models. These uncertainty requirements were designed for assimilating the albedo variable into climate, biogeochemical, hydrological, and weather forecast models. These criteria could be reasonably relaxed for other land applications, such as agriculture, land cover, land use, forestry, and hazard monitoring.

In the satellite Earth observation community, accuracy requirements for surface reflectance were first defined for the Moderate-Resolution Imaging Spectroradiometer (MODIS) product by Vermote and Kotchenova (2008), as presented in equation 1:

$$u(\rho) = 0.005 + 0.05\rho \quad 1$$

where ρ represents the surface reflectance and $u(\rho)$ its associated uncertainty. These requirements resulted from a theoretical estimation of the uncertainty budget within the MODIS

atmospheric correction process. The assessment considered uncertainties in the radiometric calibration as well as in the input parameters: atmospheric pressure, water vapour and ozone content, retrieved aerosol optical thickness (AOT) and the selected aerosol model. The same specifications were adopted in the ACIX initiative (Doxani et al., 2018) and used in the product uncertainty requirements of Sentinel-2 (Clerc et al., 2019). Consequently, they are routinely applied in the operational validation of common optical sensors like Sentinel-2 and Landsat 8/9.

2 DEFINITIONS

To provide clarity with respect to the recommendations set out in this document, it is necessary to provide definitions of key concepts and quantities that will be used throughout.

Surface reflectance is itself an ambiguous term since there are several definitions describing different types of reflectance. When used in the context of satellite products, “surface” is used as a means of describing a correction for the effects of atmospheric reflectance, transmittance, and absorption processes on the signal reflected from objects on the ground. Even in this case, there is ambiguity in whether this means that the atmospheric effects have been removed as in 1) no atmosphere, or 2) at the bottom of the atmosphere. With this said, surface reflectance is still useful in describing a group of quantities and should be used in this context. It is not applicable when describing a single physical quantity, satellite product (unless multiple quantities are packaged in one), or measurand. The nine reflectance and nine reflectance factor combinations are defined explicitly in Nicodemus et al., (1977) and Schaepman-Strub et al., (2006) and the reader is referred to those texts for further reading and the mathematical formulations. Here we provide a brief overview necessary for the context of this document; a graphical presentation of the geometries can be found in Figure 1. Before describing the various reflectance quantities there are some measurands that need to be outlined first:

Radiance: the amount of power contained in a beam per unit area and solid angle ($\text{W m}^{-2} \text{sr}^{-1}$). The distribution of radiance as a function of wavelength is called the **spectral radiance** ($\text{W m}^{-2} \text{sr}^{-1} \text{nm}^{-1}$).

Irradiance: the amount of power reaching a surface from all angles per unit area (W m^{-2}). The distribution of irradiance as a function of wavelength is called the **spectral irradiance** ($\text{W m}^{-2} \text{nm}^{-1}$).

When dividing the radiance integrated angularly over the full hemisphere by the irradiance, a quantity called reflectance (denoted by ρ), albedo, or **bi-hemispherical reflectance (BHR)** is created. In the case where the sky irradiance is perfectly diffuse, this produces the white sky albedo (Lucht et al., 2000; Schaepman-Strub et al., 2006). In general, the bi-hemispherical reflectance term is considered the most specific and therefore the clearest. The so-called black sky albedo or **directional-hemispherical reflectance (DHR)** is created when the illumination is from a single point source or direction. This quantity is given as part of the MODIS product suite. The linear combination of the black and white sky albedos approximates the ambient conditions (so-called blue sky albedo, Lewis & Barnsley 1994; Lucht et al., 2000, Schaepman-Strub et al., 2006).

The concept of a **reflectance factor** (denoted by R) is given by the ratio of the reflected radiance from a target with the reflected radiance from a Lambertian target at the same illumination and

observation geometry. A **Lambertian target** is defined as a perfectly isotropic scatterer, meaning that the reflected energy is scattered from the target equally at all angles. Lambertian targets are approximated in the laboratory and field by **reference panels** (sometimes referred to as plaques) which are highly isotropic (but not perfectly so). As a result, there will be differences between the reflectance factors derived from these targets compared to the true, unknowable quantity. Lambertian reference panels are available from several manufacturers at different nominal reflectance values (note that care needs to be taken as to the calibrated quantity – for example, manufacturers may calibrate the reflectance of the panel using a direct beam or diffuse source). In this document, a field-deployed Lambertian panel (or reference panel) can refer to any target where its primary purpose is to calculate reflectance factors through calibration. Within this there are several notable reflectance factors:

The **biconical reflectance factor (BCR)** is produced when the radiance onto and reflected from the target is defined by cones of given solid angles. The general form of the reflectance where the cones become infinitesimally small on both sides is the **bidirectional reflectance factor (BRF)**. When the incoming cone reaches its maximum solid angle, the reflectance quantity produced is the **hemispherical-conical reflectance factor (HCRF)**, which is typical of natural field conditions. If the radiance reflected from the target can be considered uniform across the cone, then this quantity is equivalent to the **hemispherical-directional reflectance factor (HDRF)**. Finally, the **bidirectional reflectance distribution function (BRDF)** represents the intrinsic scattering properties of the surface. It describes the underlying relationship between the reflectance at each incoming and outgoing geometrical configuration. From this, all reflectance and reflectance factor quantities can be derived.


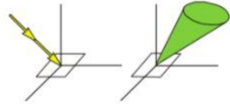

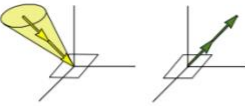
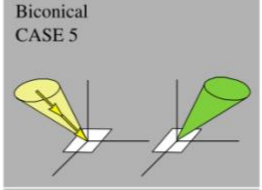
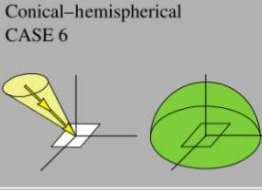

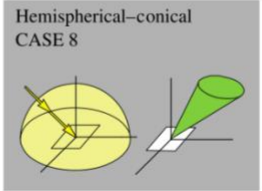
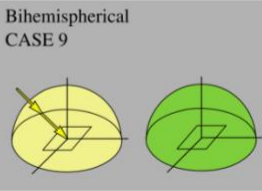
Incoming/Reflected	Directional	Conical	Hemispherical
<i>Directional</i>	Bidirectional CASE 1 	Directional–conical CASE 2 	Directional–hemispherical CASE 3 
<i>Conical</i>	Conical–directional CASE 4 	Biconical CASE 5 	Conical–hemispherical CASE 6 
<i>Hemispherical</i>	Hemispherical–directional CASE 7 	Hemispherical–conical CASE 8 	Bihemispherical CASE 9 

Figure 1. Reflectance and reflectance factor options for incoming and outgoing flux configurations; figure replicated from Table 2 of Schaepman-Strub et al., (2006). The grey boxes indicate conceptually measurable quantities; cases 8 and 9 are the only measurable quantities in the field.

The cones described in these definitions can refer to the observation or illumination geometry. For the observation geometry, this is defined by the **angular field of view (AFOV)** of the sensor which observes at a given zenith and azimuth angle. The area observed on the ground is its **field of view (FOV)** or **ground sampling distance (GSD)** and has units of distance with the same assumption applied. However, in the literature, it is common for the AFOV to be referred to as the FOV; in this document, we refer to the convention described above.

Other important terms used in this document are:

Calibration: an operation performed on a measuring system that: 1) establishes a relationship between the values given by that measurement system and values and uncertainties from a measurement standard, and 2) uses this information to convert between the value provided by the measurement system and the units given by the measurement standard (JCGM 2021: pp. 39).

Uncertainty: a non-negative parameter describing the dispersion of values being attributed to a measurand (JCGM 2021: pp. 20). The reported uncertainty must include the contribution of all non-negligible uncertainty contributions (e.g. input parameters, definitions, assumptions, etc.) propagated through to the final uncertainty.

Traceability: the property of a measurement result whereby the result can be related to a reference through a documented unbroken chain of calibrations, each contributing to the measurement uncertainty (JCGM 2021: pp. 40).

3 GENERAL CONSIDERATION FOR SATELLITE SR PRODUCT VALIDATION

Validation of satellite products is intended to ensure that the final product meets a set of requirements (e.g. product requirements, ECV requirements, etc.). In the case of surface reflectance, the definition of the final product may also include supporting products like classification maps, cloud masking, vertical column aerosol optical thickness, vertical column water vapour, and water-leaving reflectance, since they are mission dependent. The validation of satellite SR products requires a few factors related to the satellite products to be considered.

3.1 Quantities

The first of these is the reflectance quantity of the satellite product, which is not always mentioned explicitly. If this is the case, then investigation into the retrieval algorithm should be conducted, focusing on the atmospheric correction algorithm. In particular, the user should ascertain whether the diffuse sky irradiance has been removed, which would produce biconical reflectance factor (BCRF) or bidirectional reflectance factor (BRF) if assumed to be isotropic across the respective cones. Similarly, for some products (MCD43 - MODIS) the reflectance is corrected to nadir-view geometry and so doesn't reflect the original acquisition geometry of the satellite sensor. In the absence of processing designed to retrieve other reflectance quantities, the hemispherical-conical reflectance factor (HCRF) is produced (or hemispherical-directional reflectance factor – HDRF – if the reflectance is the same across the viewing cone). Deviations between HCRF/HDRF and BRF/BCRF have been shown to be large in simulation experiments (Schunke et al., 2023).

3.2 Surface

The characteristics of the surface being targeted significantly influence the design of a validation experiment. The size/extent of the site is important since it must be larger than a single satellite pixel (typically 3 x 3 pixels is the general recommendation). Larger sites require greater sampling and consequently take longer to conduct. Similarly, the spatial variability of the site (typically referred to as homogeneity or heterogeneity) determines the sampling scheme, with larger surface variability (increasing heterogeneity) requiring greater sampling. Consideration must also be taken of the surrounding areas since adjacency effects can impact the results of the validation.

The site constituents also determine the methods that can be used to measure them. For example, when conducting ground-based sampling, many sites will be disturbed by the presence of the operator and therefore do not correspond with their state when left undisturbed (Figure 2). Examples of these types of surfaces include water, desert sand, and low/grassy vegetation. Conversely, some surfaces such as tall forest canopies are inaccessible for operators and manual sampling is not practical.

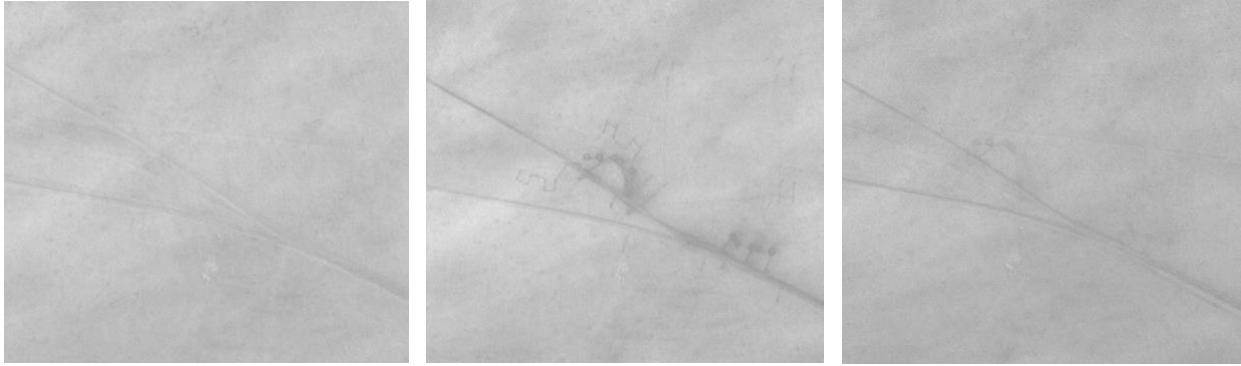


Figure 2. PLEIADES 70 cm panchromatic imagery taken before (left), shortly after (middle), and ~5 months after (right) a field campaign in Gobabeb, Namibia where manually operated spectrometers were used. Dark lines present in the central image show the impact of foot-traffic on the overall site reflectance. Image from Bialek et al., (2016) & RadCalNet Technical Working Group (2018) (copyright CNES, Distribution Airbus Defence and Space).

Another consideration is the potential change that the surface could exhibit through time (e.g. between the ground measurements and the satellite overpass). Such changes could be brought about by precipitation, drought, contamination (e.g. with dust), and vegetation phenology.

3.3 Anisotropy

Some surfaces and targets will exhibit different reflectance properties when measured from different directions. It is important to know the anisotropy of the target being measured when determining the configuration of the instruments used on the ground. For example, for a target with a strong BRDF, the ground instrumentation should measure in the same angular configuration as the satellite sensor to ensure equal comparison.

3.4 Illumination considerations

As well as the target changing, the illumination conditions will also change in time in relation to the solar zenith/azimuth angle or potentially cloud cover. The solar zenith angle change is of particular importance since the reflectance of anisotropic surfaces will change in response (Figure 3). Assessing the maximum acceptable change in solar zenith angle will subsequently determine the maximum permissible difference in time between the satellite and in situ observations (Origo et al., 2020). A second factor to consider is the cloud cover present; ideally validation measurements should be taken during clear sky conditions to avoid cloud shadowing of the surface and increased contribution to the diffuse sky irradiance.

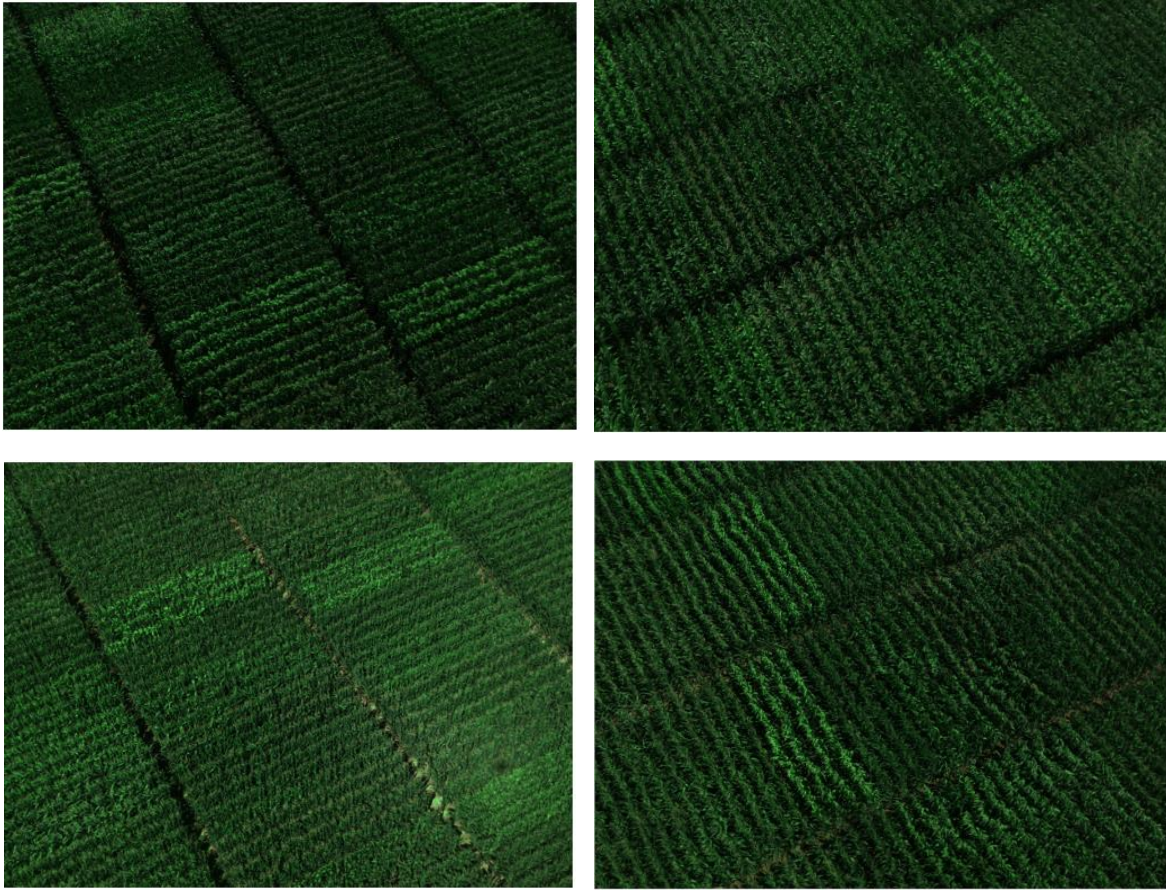


Figure 3. Four images taken from different view angles of the same location over the corn site used for the SRIX4Veg campaign in Las Tiesas, Barrax, Spain. Clockwise from top left: 45°, 135°, 225°, and 315° view azimuth angles. The bottom left image shows the highest brightness since it is closest to the hotspot. Images courtesy of Daniele Latini (Geo-K).

3.5 Spectral considerations

The spectral reflectance of a target will likely vary with wavelength because of its absorption, transmittance, and reflectance properties. This means that measurements of the surface must be spectrally equivalent to the satellite sensor being validated. This can be done by using a broadband filter with a similar spectral response or by subsampling the spectral (ir)radiance at the surface, convolving it with the satellite sensor spectral response function, and calculating the resulting band reflectance. For fine satellite spectral bands, it may be sufficient to assume that the variation in spectral reflectance across that band is sufficiently small as to be considered negligible and therefore not warrant matching of the spectral response.

4 GENERAL STRATEGIES FOR COLLECTION OF SR REFERENCE DATASETS

There are many ways of collecting surface reflectance data suitable for validating EO products and the area has been built from decades of experience in field spectrometry across the globe (Goetz 1975; Goetz 1987; Teillet et al., 2001; Campbell et al., 2001; Schaepman-Strub et al., 2006; Milton et al., 2009; Anderson et al., 2012; Hueni et al., 2017; Doelling et al., 2017; Malthus et al., 2019a). The evolution of this area has come from the advancement of instrumentation, increasing focus on the fidelity of the measurement procedure, and the need for accurate uncertainty estimates. The field spectrometer remains a key piece of instrumentation for surface reflectance measurements. Early field spectrometers, developed for the validation of Landsat satellites, were primarily research instruments and were heavy and cumbersome to operate (Goetz 1975; 1987). The utility and popularity of these instruments has driven progressive development of more agile and powerful field instruments over the last few decades (Goetz 2009). The weight of commercially available field spectrometers today ranges from 5 kg – 7 kg, down from the 30 kg instrument used in Goetz (1987). Similarly, the time to acquire spectra has decreased hugely from 30 s in Goetz (1975) to the present-day range which is under 1 s.

Understanding of these instruments and their measurements has also gradually improved. For example, the effect of detector jumps (i.e. step changes between spectral regions measured by different detectors) was highlighted as a major issue in Milton et al., (2009) such that they are present in reflectance spectrums. Hueni & Bialek (2017) derived a physically based explanation for this effect based on external temperatures and produced a correction method which reduced these errors to a maximum of 0.6%. Similarly, advances in the calibration and characterisation of reference targets have meant that derivation of the in-field hemispherical-conical reflectance factor (HCRF) can be based on tailored calibration coefficients that are normalised to the solar geometry and the proportion of diffuse and direct sky irradiance (Origo et al., 2020). The authors report that the reference panel deviation from the ideal Lambertian target can be as high as 3.8%.

The potential for pitfalls associated with making spectral reflectance measurements in the field has produced several good practice protocols and recommendation works designed to standardise the way these measurements are made for specific applications. For example, Schweiger (2020) provides recommendations on the acquisition of spectra for plant biodiversity, Gras et al., (2014) provide guidance on the collection of spectral reflectance for soils, and Behnert et al., (2011) tailor the instructions to the characterisation of land test sites. Similarly, instructional materials and protocols on making spectral measurements with specific instruments can be found in Kalacska et al., (2019) and the UK's National Environmental Research Council (NERC) Field Spectroscopy Facility (FSF) resources (<https://fsf.nerc.ac.uk/resources/guides/> [link accessed: 24/02/2024]), while general purpose EO validation and field data acquisition recommendations can be found in Soto-Berelov et al., (2015) and Trevithick (2015). A detailed example is provided in the case study at the end of this chapter.

The field spectrometer offers a means of measuring reflectance spatially (through sampling) and temporally through repeat measurements. For satellite validation, the scales of operation in both dimensions have a maximum of a few hundred metres spatially to several hours temporally. However, the exact values are co-dependent with greater spatial coverage reducing the temporal frequency, and vice versa. This means that field spectrometers are restricted to homogeneous areas and subsampling (as in Malthus et al., 2019a; Origo et al., 2020).

A key improvement in the temporal domain came with the advent of autonomous field spectrometers such as the Multi-Angle Spectrometer (MAS) which was a tower-deployable system capable of observing surface radiance from multiple view azimuth angles as well as sky irradiance and dark current (Leuning et al., 2006). This and other novel systems formed part of the Spectral Network (SpecNet) which featured an international network of cooperating investigators (Gamon et al., 2006a). However, it wasn't until the LANDHYPERNETS network that an autonomous spectrometer system tailored to satellite validation of the land surface was produced. This featured a common processing system (Goyens et al., 2021) designed to ensure traceability and uncertainty propagation to the spectral reflectance produced. The system is capable of measuring at multiple view angles to provide data for various satellite sensors overpassing a particular site. A demonstrator of this was recently published in Morris et al., (2024). An equivalent network, known as WATERHYPERNETS, features the same philosophy (with slight instrumental differences) over water sites (Doxaran et al., 2023; González Vilas et al., 2024) and was based on the earlier Pan-and-Tilt Hyperspectral Radiometer System (PANTHYR; Vansteenwegen et al., 2019). The single-sensor fixed-view nature of these instruments means that the site locations must be homogeneous in space over the footprint of the targeted satellites and have a spatially uniform BRDF (since the spectrometer views from in-to-out rather than out-to-in as in the definition of BRDF).

Methods for sampling the spatial domain have been varied with different platforms ranging from tracks (Berry et al., 1978; Brach et al., 1983; Gamon et al., 2006b), tractors (Steven et al., 2004), bicycles (Milton et al., 2009 fig 2e), balloons (Chen & Vierling 2006), and aircrafts such as planes and helicopters (Milton et al., 1994; Milton et al., 2009), with most featuring single pixel spectrometers. For platforms which are at ground level (bicycles, tractors, trams/tracks, etc.), there are two issues: 1) the platform actively disturbs the surface and/or 2) there is the potential for shadowing. For those which are aerial in nature (balloons and aircraft), there is a large cost associated with each measurement campaign (due to hiring pilots and fitting instrumentation in aircraft) or the motion of the platform is not controllable (e.g. balloons, kites, etc.) and as such the area being measured on the ground is difficult to predict.

UAV platforms offer a convenient middle ground between the aforementioned techniques since they are relatively cheap, require limited training (depending on the country but approximately 10 hours of practice), are programmable for autonomous flights (at the current time with a pilot observing), have good repeatability in terms of geolocation, are small enough that shadowing is not generally an issue, and do not disturb the surface. Compact hyperspectral imagers are also

now available on the market which can produce hyperspectral data cubes in the 400 nm – 2500 nm wavelength range and are tailored to UAV-platforms. This means they can provide wall-to-wall sampling of an area of approximately 100 m x 100 m or larger (Arroyo-Mora et al., 2023), which is much greater than other techniques.

4.1 Case study: CEOS-WGCV-endorsed surface reflectance protocol for ground-based measurements

The CEOS WGCV endorsed protocol has been designed to support the systematic collection of in situ field spectroscopic data for the purposes of validation of surface reflectance satellite products across large regional scales such as the continent of Australia. The purpose of the protocol is to provide guidelines on the collection of data to ensure a high level of consistency between teams, including the selection of specific sites for measurement and in the systematic collection of both site description data and of ground-level spectral radiance and reflectance measurements using field spectroradiometers or spectrometers. The methods are designed to be pragmatic, straightforward, and repeatable, and should ensure consistency and enduring value in the data collected.

To enable the verification of satellite-derived reflectance products, conventional ground-based observations are required and should be collected using calibrated and traceable field instrumentation and associated methods. It is only when the uncertainty of the data products is quantified that can we then develop improved algorithms to reduce the uncertainty (Malthus et al., 2014). This process seeks understanding of the different types of uncertainty and their dependency on the overall uncertainty. For example, where an uncertainty is statistically derived it may be feasible to reduce the uncertainty by undertaking more measurements.

The delivery of accurate and precise surface radiance and reflectance data through the application of a repeatable field protocol is the goal, the rationale for which is illustrated in Figure 4. Validation data is acquired by field teams through regular visits to the field sites at the time of satellite overpasses. Methods can be radiance-based (Milton et al., 2009) or reflectance-based (Thome, 2001, Thome et al., 2005). This protocol addresses the radiance-based method.

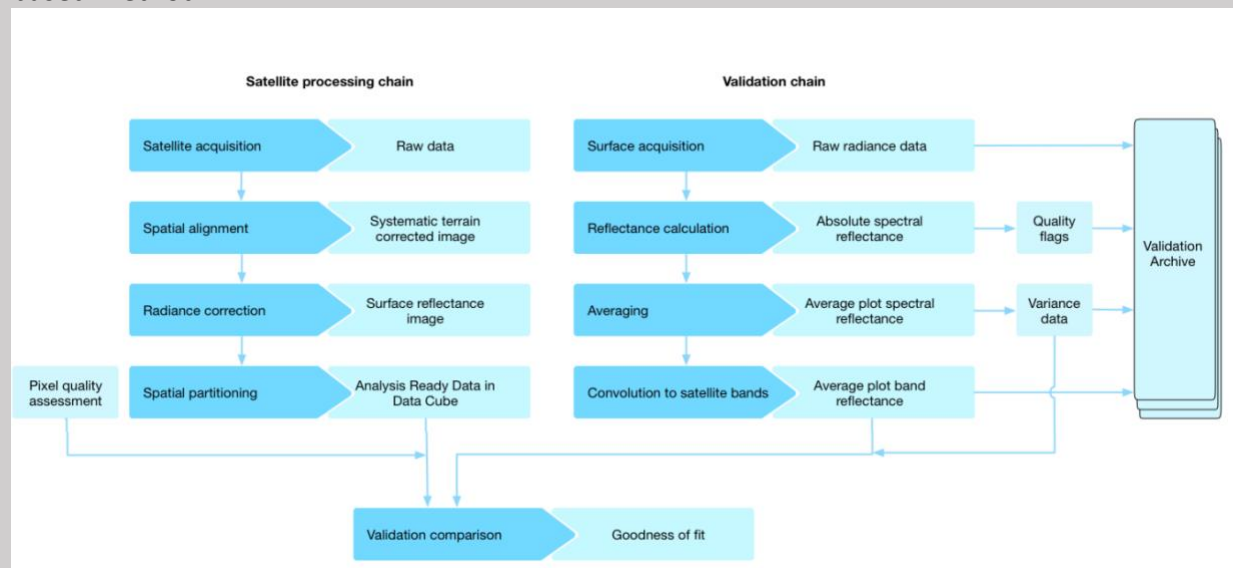


Figure 4 Diagram illustrating the role that field measurements play in the validation of Analysis Ready Data. The left side of the diagram illustrates the satellite data process flow from raw data to Analysis Ready Data (summarised from Lewis et al., 2017). The right side shows the field data acquisition and data process flow.

The ground-based surface reflectance protocol covers a wide variety of topics related to the data collection. In terms of the timing of the data collection, it discusses the requirements for deciding on the go/no-go of the campaign, such as cloud cover and rainfall, and the best resources for determining the satellite overpass time. In terms of the method of data collection, a comprehensive list of essential equipment is provided, treatment of that equipment (e.g. warm-up times, setup instructions, etc.), transect setup, measurement instructions (e.g. optimal configurations of instruments, operational instructions, etc.), and data recording (including ancillary information). The roles of each operator are discussed, with the primary operator (the person using the spectrometer) focused on the operation of the spectrometer, while the secondary operator is focused on recording the conditions at the time of the measurement, noting anomalous incidents, keeping track of scan numbers, and monitoring the primary operator's actions to spot mistakes.

Figure 5 gives a schematic example of the setup of a 100 m x 100 m site to facilitate the transect sampling design. The whole site is orientated to the satellite overpass angle and a reference panel is placed for easy optimisation at the end of each line. For further information the reader is referred to Malthus et al., (2019a).

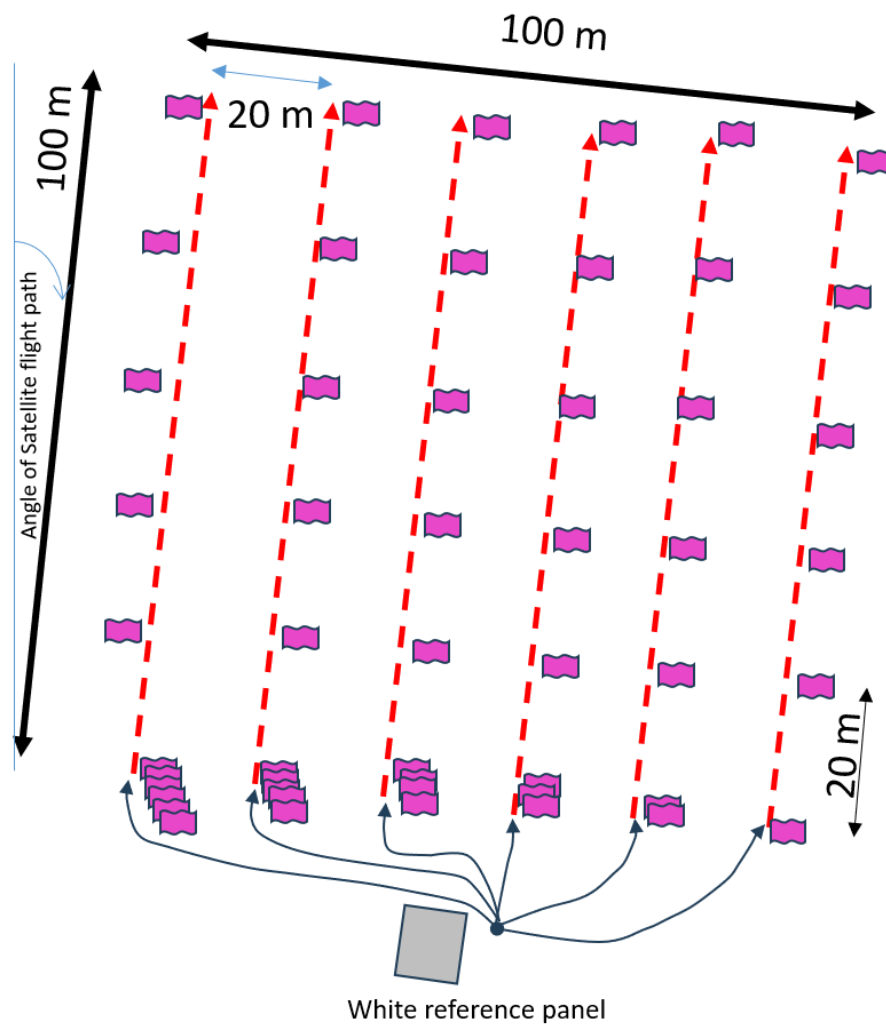


Figure 5. Proposed field sampling design of a 100 m x 100 m area consisting of 6 transects 20 m apart, acquired in the Southern Hemisphere, where the satellite flight path is 8 degrees from true North. The spectroradiometer operator walks in a northerly direction after first measuring the white reference panel and returns to the panel along the same path to collect a second panel measurement before repeating the process. The pink polygons represent the flags marking every 20 m, at the end of each line the number of flags represents the line number. The white reference panel is given by the grey square.

5 GOOD PRACTICES FOR UAV-BASED SR REFERENCES (CONDENSED VERSION)

This section provides a condensed version of the good practice recommendations and requirements set out in chapter 6. Requirements are given for activities aiming to meet CEOS-FRM standards (see Goryl et al., 2023 for more information), where producing CEOS-FRM datasets is not the goal these requirements can be relaxed to recommendations. Users are referred to this chapter for greater detail and context on the recommendations. The recommendations and requirements are split into subsections, as such the recommendations do not appear in the order they appear in section 6:

5.1 Sensor

[R1] As a minimum, we *recommend* the use of spectrometers capable of performing contiguous measurements at a spectral resolution no coarser than 10 nm.

[R20] Regular calibration and characterisation, between campaign seasons, is *required* to fully understand the uncertainties pertaining to all instruments and artefacts.

[R21] Many systems will include an automatic dark collection immediately prior to, or following, the exposed target measurements (ideal). If this is not present, *it is required* that manual dark readings should be performed before and after the flight.

5.2 UAS

[R2] At the present time rotary-wing UAVs are *recommended* for spectroscopic applications owing to their manoeuvrability and precise control over flying speed, which is particularly important when acquiring data at low flying heights with pushbroom imagers.

[R3] The use of gimballed mounting is *recommended* to compensate for in-flight pitch, roll, and yaw movements of the platform.

[R4] We *recommend* using platforms with in-flight real-time kinematic (RTK) or differential global positioning system (DGPS) positioning as they offer more accurate positioning and higher in-flight stability.

[R7] A co-aligned LiDAR system is a *recommended* solution for pushbroom scanning, though it requires significant trade-offs since LiDAR utilisation increases the weight and size of the payload, as well as the overall energy consumption, effectively shortening achievable flight duration [R7]. If these trade-offs are not acceptable, acquiring the LiDAR data separately to the imaging flight will allow for a longer flight time.

5.3 Environmental conditions

[R13] Clear sky conditions are *recommended* for high quality surface reflectance validation activities. Ideally the sky should be completely clear of cloud and haze for the highest quality data, however, small amounts of cloud could be tolerated if it is well away from the solar disc and low in the sky.

5.4 Flight planning

[R14] *It is recommended* to set the orientation of the flight lines to be parallel with the overpass angle of the satellite being validated.

[R15] Developing the flight lines using Geographic Information System (GIS) software is *highly recommended*.

[R17] A minimum central area of 9 x 9 pixels from which to extract mean and standard deviation with a minimum perimeter of 3 pixels, to eliminate any edge effects due to spectrometer point spread function and scattered light issues, *is recommended*.

5.5 Supplementary data

[R16] For the reference targets, a minimum of two radiometrically distinct (different reflectance levels) targets *are required*; one should be below the lowest HCRF expected in the field and one should be above that.

[R5] Supplementary GNSS/IMU systems employed with the sensor can contribute to improved positional accuracy and are *recommended*.

[R6] It is *recommended* to first assess the georeferencing accuracy under different conditions, especially when relying on manufacturer-provided workflows, to establish whether ground control points (GCPs) for absolute positioning would still be required.

[R8] For all GNSS processing, the use of PPK or RTK based on high-quality reference data (e.g. CORS network) is *recommended*.

[R9] It is *recommended* that the user determine the georeferencing requirements based on the spatial resolution of the satellite product:

- >10 m spatial resolution: direct georeferencing with high-quality GNSS/IMU systems is sufficient.
- 5 m – 10 m: at least geometry evaluation with GCPs should be considered after direct georeferencing.
- <5 m resolution: georeferencing with GCPs is recommended after direct georeferencing to guarantee geometric quality.

[R10] For the highest accuracy we would *recommend* the use of a locally established GNSS base station, or otherwise one located no further than 10 km from the study site.

[R11] For UAVs, in areas with a relatively unobstructed sky view we *recommend* opting for RTK positioning that allows instantaneous collection of positions at centimetre-level accuracy; this solution requires cellular communication and real-time access to a network of reference GNSS stations.

[R12] In difficult terrain, we *recommend* the use of multi-band GNSS receivers that are capable of measuring across several frequency bands increasing the chances of obtaining a fix.

5.6 Post-processing

[R18] It is *recommended* that view zenith angle filtering is applied even if it results in data gaps (assuming the target area is reasonably homogeneous).

[R19] It is *recommended* that hyperspectral data is orthorectified but not converted into a raster format (i.e. gridding or stitching).

[R22] *It is required* that the uncertainty of all processed data is provided with the results and that the calculation of that uncertainty follows the guidance in the “Uncertainty characterisation and propagation” section.

[R23] The conformity analysis method of validation is *required* where adherence to the FRM guidelines according to Goryl et al., (2023) is needed [R23].

6 GOOD PRACTICES FOR UAV-BASED SR REFERENCES

This chapter aims to bring together the current state-of-the-art of UAV-based hyperspectral radiometry for surface reflectance product validation for the purpose of providing readers with sensor-agnostic good practice guidance. It is recognised that since this is a new field of research there will be gaps in the community's knowledge; where possible, these are identified and discussed. Likewise, the guidelines are discussed with respect to the technology available at the present time of writing (November 2024) and it is intended that these guidelines will be updated in future editions of this document as scientific understanding and technological capabilities progress. This section is partitioned into subsections which broadly discuss the measurement and processing techniques required to produce reflectance quantities which match those produced by the satellite products for the purpose of validation.

6.1 Reflectance methods

The downward-facing optical sensors mounted on UAVs typically measure at-sensor radiance in digital numbers. Radiance alone is not sufficient to estimate the reflectance of the target in the scene. As a result, two main methods exist for retrieving surface reflectance from UAV-mounted sensors: irradiance- or target-based. However, within both methods there are many variations (further information can be found in Aasen et al., 2018).

The most common method at the present is the **target-based approach** which utilises reference targets, this is commonly known as the empirical line method (ELM; Honkavaara & Khromashahi 2018). In this scenario, the UAV-mounted sensor retrieves the reflectance factor of the scene by comparing the signal measured over the reference panel (of known reflectance) and the rest of the scene. The reference panel material is designed to approximate a Lambertian scatterer, which is required in the estimate of reflectance factor quantities. The practicalities associated with the number of reference targets, materials, reflectance levels, and the equations associated with the different approaches is discussed later in this section.

The **irradiance-based approach** uses an estimate of the at-sensor irradiance as the reference from which the reflectance can be calculated. The source of irradiance may be derived from a dedicated sensor or from atmospheric radiative transfer models (e.g. Zarco-Tejada et al., 2012; note that if this method is used then the extraterrestrial solar irradiance spectrum used should be specified) forced by measurements of atmospheric properties such as aerosol optical depth and water vapour (Aasen et al., 2018). The dedicated sensor approach is recommended if using this method because, through calibration of the sensor, a link to SI is established; it also makes rigorous characterisation of the uncertainties easier. Irradiance-based methods can use spectrometers onboard the UAV (as per Nevalainen et al., 2017 & Suomalainen et al., 2021), however, the current weight limitations mean that this approach is primarily limited to multispectral instruments. Levelling of the irradiance foreoptic is still a challenge despite novel methods to accomplish this (e.g. Suomalainen et al., 2018), since large differences in the irradiance can be produced when levelling is not sufficient (Burkhart et al., 2017). Alternatively,

spectrometers on the ground can be used to provide the irradiance from a single location, thereby removing the weight limitations from the UAV platform and allowing for greater control over the levelling of the foreoptic (Suomalainen et al., 2021). This approach typically relies on the illumination conditions being representative between the UAV and the irradiance sensor on the ground. The irradiance-based approach offers a dynamic estimation of the reflectance assuming that the two measurements are conducted at the same time. This is particularly attractive for many UAV applications. However, for satellite validation purposes a stable sky irradiance is required for the duration of the measurement because the satellite imaging is instantaneous. Therefore, this approach would serve as an additional means of assessing that uncertainty rather than meaning that measurements could be taken in a wider variety of conditions.

6.2 Instrument types and supplementary UAV system components

UAV-mounted spectrometers fall into three main categories: (i) point spectrometers, (ii) pushbroom line scanners, and (iii) 2D camera-based imaging systems (Figure 6). Point spectrometers provide spectral signatures over the sensor's footprint which is dictated by the AFOV and the flying height. They offer fine spectral resolution, high dynamic range, and high signal-to-noise ratio (SNR). This comes at the expense of the spatial dimension since only selected locations are being sampled.



Figure 6. Clockwise for top left: Low Altitude Advanced Hyperspectral System (LAHYS) developed by NRC and the Applied Remote Sensing Laboratory at McGill University, which features a HySpex pushbroom imager (image credit: Juan Pablo Arroyo-Mora; Arroyo-Mora et al., (2023)); Rikola 2D frame hyperspectral imager from the National Land Survey, Finland (image credit: Eija Honkavaara); Ocean Optics STS point spectrometer from Maitec (image credit: Stefan Maier); Headwall Co-aligned pushbroom imaging spectrometer at NPL (image credit: Rosalinda Morrone).

For mapping applications, imaging spectrometers are typically chosen as they facilitate capturing spatially resolved spectral responses albeit not without some challenges in fully understanding the spatial coverage of the resulting imagery (Inamdar et al., 2023). Pushbroom scanning, whereby the area is scanned line-by-line in-flight to generate a continuous image, is commonly employed in such spectrometers. It offers high spatial and spectral resolution but typically requires robust external GNSS/IMU and digital surface model (DSM) information to enable accurate georeferencing and orthorectification. In the case of 2D camera-based imaging systems, which instantaneously record spectral information in two spatial dimensions, spectral and 3D information can be obtained from an instantaneously acquired frame. This is thanks to the rigid imaging geometry that allows employment of the Structure from Motion (SfM) workflow and subsequent DSM creation.

The selection of the most appropriate sensor will be dictated by the required spectral and spatial resolution; the trade-offs between different imaging approaches can be further explored in Aasen et al., (2018). As a minimum, it is *recommended* that spectrometers capable of performing contiguous measurements at a spectral resolution no coarser than 10 nm are utilised to collect data for SR product validation [R1]. This is because most multispectral satellite instruments do not feature spectral bands narrower than this. For hyperspectral satellite missions, finer spectral bands on the UAV-mounted instrument will be required.

The requirements imposed by the utilised sensor will often dictate platform selection with the main restricting factors being the sensor's physical dimensions and weight. Although fixed-wing platforms facilitate larger area coverage, at the present time rotary-wing UAVs are *recommended* for spectroscopic applications owing to their manoeuvrability and precise control over flying speed, which is particularly important when acquiring data at low flying heights with pushbroom imagers [R2].

The in-flight stability of the sensor ensuring consistent viewing direction is one of the key criteria to consider. A hard mount offers an alternative where the use of a gimbal is not possible due to weight limitations. However, data quality may be impacted since the effective view zenith angle will be influenced by the platform movements – this can prove problematic in windy conditions. The use of gimballed mounting is *recommended* to compensate for in-flight pitch, roll, and yaw movements of the platform [R3]. Platforms utilising in-flight real-time kinematic (RTK) or differential global positioning system (DGPS) positioning can offer further advantages, allowing more accurate positioning – their use is *recommended* as they ensure better in-flight stability [R4].

Supplementary GNSS/IMU systems employed with the sensor *are recommended* for improved data quality [R5] with the quality of the respective system related to the positional uncertainty; the quality of the GNSS/IMU system should be based on the uncertainty requirements of the application. Properly bundled (calibrated) GNSS data provides precise location of the sensor aperture, whilst the IMU measures sensor orientation and motion. This combination not only allows for direct georeferencing but also for determination of the effective view zenith angle for

each pixel within an image. There are a range of available GNSS/IMU systems on the market, which offer different accuracy levels. The quality of georeferencing and orthorectification will additionally depend on the distance to a base GNSS station against which recorded UAV flight trajectories are being corrected, and the deployed processing workflow. It is *recommended* to first assess the georeferencing accuracy under different conditions, especially when relying on manufacturer-provided workflows, to establish whether GCPs for absolute positioning would still be required [R6] to achieve the desired geolocation accuracy.

Accurate georeferencing is important for orthorectifying data from hyperspectral sensors mounted on UAVs. Accurate georeferencing reduces the number of UAV-instrument pixels which are wrongly allocated to the satellite pixel of choice. The use of external DSMs necessitates good data co-registration to prevent the introduction of processing artifacts. An alternative is to employ a co-aligned LiDAR system that allows creation of local and detailed DSMs. Such LiDAR point clouds are typically georeferenced utilising trajectory readings from the same GNSS/IMU system, solving the issue of data misalignment. A co-aligned LiDAR system is a *recommended* solution for pushbroom scanning, though it requires significant trade-offs since LiDAR utilisation increases the weight and size of the payload, as well as the overall energy consumption, effectively shortening achievable flight duration [R7]. If these trade-offs are not acceptable, acquiring the LiDAR data separately to the imaging flight will allow for a longer flight time.

6.3 Geometry

The geometric accuracy of UAV spectral products refers to the spatial accuracy with respect to a global reference frame, for example a geodetic datum like WGS84 and a coordinate reference system like UTM. This accuracy is relevant to satellite validation because the validation product needs to spatially coincide with the satellite product. Only spatially accurate UAV products can guarantee the validation of the same land surface as the satellite pixel. As UAV sensors are built to produce very high spatial resolution products and geometry can be accurately controlled, the geometric uncertainty of UAV hyperspectral reflectance can be minimised with minimal effort, to the level where it can be neglected in the validation process.

In general, there are two strategies for georeferencing (UAV) spatial products: direct and indirect georeferencing. Direct georeferencing relies on continuous collection of GNSS and IMU data on the UAV along with the spectral data, calibration of the lever-arms between the sensor optical centre and antenna phase centre, boresight calibration to determine the camera and IMU axis misalignment, and geometric calibration of the sensor optical system. After data collection, GNSS and IMU data is processed to a UAV trajectory (i.e. position and orientation of the platform and sensor), possibly with external GNSS reference data (i.e. Post-Processing Kinematic (PPK)). This trajectory can be used to georectify the spectral data, making GCPs obsolete. Direct georeferencing of UAV data can achieve spatial accuracies at the centimetre-scale (Pfeiffer et al., 2012; Dreier et al., 2021) but its quality strongly depends on the utilised GNSS and IMU sensors, the calibration (camera, lever-arm, and boresight) and external correction data (PPK processing).

Indirect georeferencing relies on GCPs that are laid out in the field and surveyed with survey equipment like RTK-GNSS devices. After initial geometric processing of the spectral data, an absolute uncertainty estimate is achieved by matching the surveyed GCPs with their counterparts in the produced imagery. Next to the additional field efforts, the GCPs must be recognisable in the UAV products. The indirect georeferencing approach can only be applied to 2D-imagers since the position and orientation of the sensor must be fixed for the instance that an image was collected. For pushbroom imagers, the position and orientation of the sensor is changing for every new line collected, requiring employment of the direct georeferencing approach. That said, a second stage georeferencing involving GCPs can be used to improve the accuracy.

Recommendations concerning geometry:

- For all GNSS processing: use PPK or RTK based on high-quality reference data (e.g. CORS network) [R8]
- The preferred georeferencing approach depends on the spatial resolution of the satellite product [R9]:
 - >10 m spatial resolution: direct georeferencing with high-quality GNSS/IMU systems is sufficient.
 - 5 m – 10 m: at least geometry evaluation with GCPs should be considered after direct georeferencing.
 - <5 m resolution: georeferencing with GCPs is recommended after direct georeferencing to guarantee geometric quality.
 - For 2D imagers, georeferencing with GCPs is recommended.

In general, the orthorectification and georeferencing procedure should utilise the operational software provided with the UAV-based hyperspectral imager since this will likely include key parameters and models which are specific to the hyperspectral instrument. However, it is noted that at the time of writing there is a lack of uncertainty propagation in this software, leaving an open gap.

6.3.1 GNSS

The need for supporting GNSS measurement will depend on the utilised UAV system, site characteristics, and available GNSS infrastructure. For example, systems relying on direct georeferencing making use of GNSS/IMU flight trajectory information will require GNSS observations from a nearby base station. Such observations can often be directly acquired from a network of Continuously Operating Reference Stations (CORS). However, we *recommend* the use of a locally established GNSS base station instead [R10]. Where this is not possible, this requirement can often be satisfied with the use of Virtual Reference Stations (VRFs), which are created by employing a DGPS algorithm to interpolate observations from several CORS stations. The use of a VRF is often a preferred choice (compared to the direct use of CORS stations) since

positioning errors will increase with the baseline distance, reducing the overall positioning accuracy.

For systems relying on indirect georeferencing, survey-grade GNSS receivers are a necessity since obtaining accurate positions of the deployed GCPs in the target coordinate system is key for successful georeferencing of UAV products. In areas with a relatively unobstructed sky view, we *recommend* opting for RTK positioning that allows instantaneous collection of positions at centimetre-level accuracy; this solution requires cellular communication and real-time access to a network of reference GNSS stations [R11]. In more complex terrain, e.g. within forest gaps where the sky view is largely obstructed, or where mobile reception is an issue, PPK positioning should be used instead. This will require post-processing against measurements from a nearby GNSS base station. For the highest accuracy, we would recommend the use of a locally established base station, or otherwise one located no further than 10 km from the study site. 5 minutes of continuous GNSS measurement at each GCP is typically sufficient to allow PPK. Periods longer than 5 minutes may be required depending on how obstructed the sky is (both over the GCP and the base station) and how far away the base station is located since post-processing requires both receivers to have simultaneous observations from the same satellites. In difficult terrain (e.g. obstructed skies), we *recommend* the use of multi-band GNSS receivers that are capable of measuring across several frequency bands increasing the chances of obtaining a fix [R12]. The GCP coordinates can alternatively be surveyed with a total station referenced to known locations, e.g. points previously surveyed through static GNSS.

Care should be taken to ensure the right coordinate system is used when obtaining GNSS measurements (both in the case of base stations and GCPs) and subsequently using these for georeferencing. In some cases, coordinate transformation is necessary if there is a mismatch between the coordinate system of the GCPs and the UAV location tags. We, therefore, suggest verifying both the horizontal coordinate systems and the vertical datums used to ensure accurate positioning of the final products.

6.4 Flight planning

Flight planning is a complex task that requires a good understanding of the intricate relationship between the UAS flight performance, the hyperspectral sensor characteristics, and the expected radiance or reflectance outputs. These three aspects are presented with the aim of highlighting the best practices for flight planning for validation of spaceborne surface reflectance products. Ultimately, flight planning is a needs-driven approach where one determines the required characteristics of the hyperspectral instrument, then works backwards based on the sensor and flight parameters to determine the appropriate combination to achieve those characteristics. To acquire an image with similar pixel size, pixel aspect ratio, signal-to-noise ratio (SNR), and overall coverage from two different UAVs equipped with different hyperspectral imagers will require different flight plans (e.g. optimal speed, altitude, etc.). The three components that need to be considered when developing a UAV hyperspectral imager (HSI) flight plan are:

- Logistical constraints
 - Flight controller capability, batteries/fuel, environmental variables (e.g. wind, temperature, etc.), flight geography, and regulatory framework.
- Sensor characteristics
 - FOV, integration and frame time, SNR, and the number of across-track pixels.
- Target characteristics of the field site
 - Estimated reflectance and BRDF.

6.4.1 UAV flight conditions

Before considering the design of the flight plan, there are a minimum set of environmental conditions that must be met. The first condition is the wind speed; it is key that the wind speed does not exceed the operating tolerance of the UAS as this is a health and safety risk. UAS pilots should be aware of the maximum wind speeds their systems can operate in and be able to measure the ambient conditions to ensure that threshold is not exceeded. Wind speeds under the maximum tolerance of the UAS can also cause issues with the measurement process and potentially increase the uncertainties. This, along with mitigations, is discussed later in this section.

The illumination conditions for the duration of the UAS flight must be stable. This is because the satellite will view the site quasi-instantaneously and it is imperative that the UAS data acquisition takes place in conditions representative of that. Strictly speaking, a validation can occur even when a site is visible through a gap in the clouds. In practice, however, only clear sky days are viable because cloud will move across the sky and create illumination conditions that are inconsistent with the time of the satellite overpass. The addition of downwelling irradiance sensors can enable calculation of the correct reflectance at the point of measurement, even in changing illumination conditions, since they remove the dependency on reflected radiance measurements of reference targets at specific points in the flight. However, this method still can't provide reflectance for the whole site representative of the illumination conditions at the time of the satellite overpass during variable conditions. As a result, clear sky conditions are *strongly recommended* for high quality surface reflectance validation activities [R13]. Ideally, the sky should be completely clear for the highest accuracy measurements, however, small pockets of cloud could be acceptable if they are low in the sky and not blocking or near the sun. It is recognised that this dependency will increase the cost associated with logistics and may result in failed acquisitions (e.g. where the forecast predicts clear skies but this turns out not to be the case); nevertheless, it remains a recommendation due to its importance.

6.4.2 UAS flight performance

Understanding the flight performance of the UAS is a key aspect of HSI data acquisition. Current advanced sensors require a heavy lift UAV capable of accommodating the HSI payload, gimbal, and other components, resulting in a total take-off weight of between ~20 kg – 40 kg. Even for small HSI that have the radiometric quality necessary for spaceborne validation activities, when

the ancillary components are considered, the combined payload weight can approximate 2 kg – 3 kg requiring a robust UAV with a maximum take-off weight (MTOW) or maximum take-off mass (MTOM) of at least 7 kg – 10 kg. Operating UAVs in these weight classes requires careful flight planning not only for maintaining safety of operations but also for the acquisition of high quality HSI. As a result, one of the first requirements of choosing an appropriate UAV platform is whether the system is capable of following a precise flight path (horizontally and vertically) and is stable in flight with smooth acceleration/deceleration. While a gimbal can compensate for roll, pitch, and yaw variability during flight, it cannot correct for drift in the flight path, changes in speed, or imprecise altitude. The impact is especially evident in the quality of the non-geocorrected data. Figure 7 compares the original geometry (no geocorrection) of two different UAS with the same hyperspectral sensor (HySpex VS-620). The figure illustrates that performance can vary greatly between UAVs and gimbals. Figure 7A shows higher pitch and roll induced variability on the raw geometry of the imagery, compared to Figure 7B. When the flight is unstable, specific targets on the ground may be missed entirely, duplicated, or smeared in the original geometry.

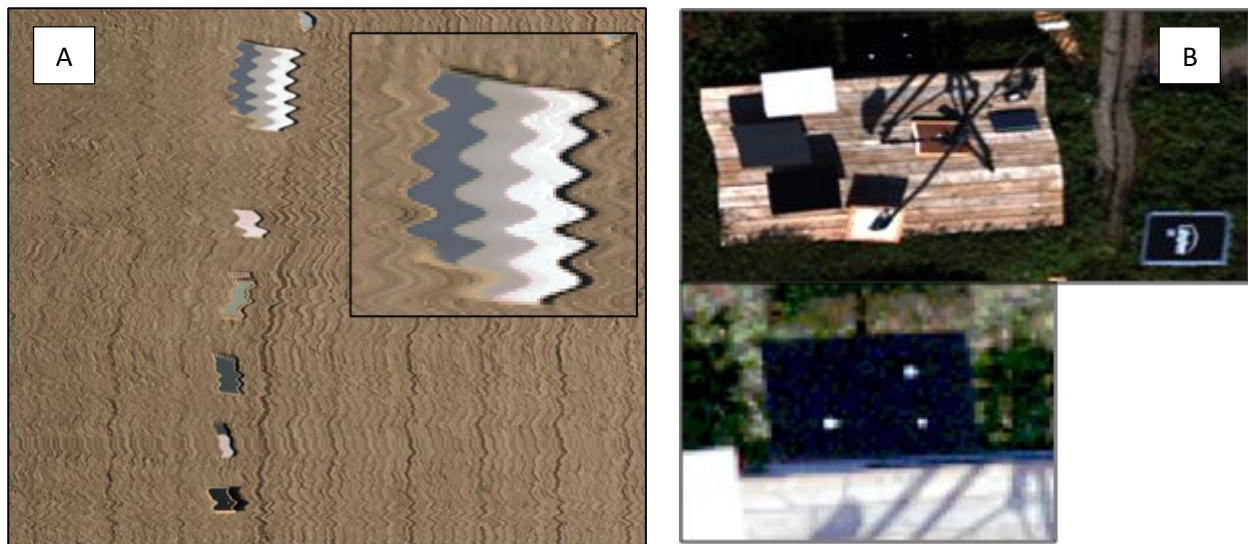


Figure 7. Example of raw imagery collected over calibration targets with two different UAS and same hyperspectral sensor – i.e. HySpex VS-620. (Left, A) HSI collected at Las Tiesas Experimental Station, Spain in July 2022. (Right, B) HSI collected at the Mer Bleue Peatland Cal/Val Supersite in July 2023. In A, pitch and roll artefacts are evident in the non-geocorrected data. In B, there are fewer spatial artefacts in the non-geocorrected data and even small targets as shown in the inset are clearly resolved.

The use of RTK with the flight controller has shown to be essential for the UAV to follow a precise flight path at the planned altitude (Arroyo-Mora et al., 2019), minimising the total system error in both the horizontal and vertical components (Figure 8). In the horizontal (planar) component, RTK allows for reducing the flight technical error and navigation solution error. In the vertical component RTK also reduces the altimetry error. With an RTK-enabled flight controller, these contributors to the total system error may be reduced to the centimetre scale as opposed to decimetre or metre scale.

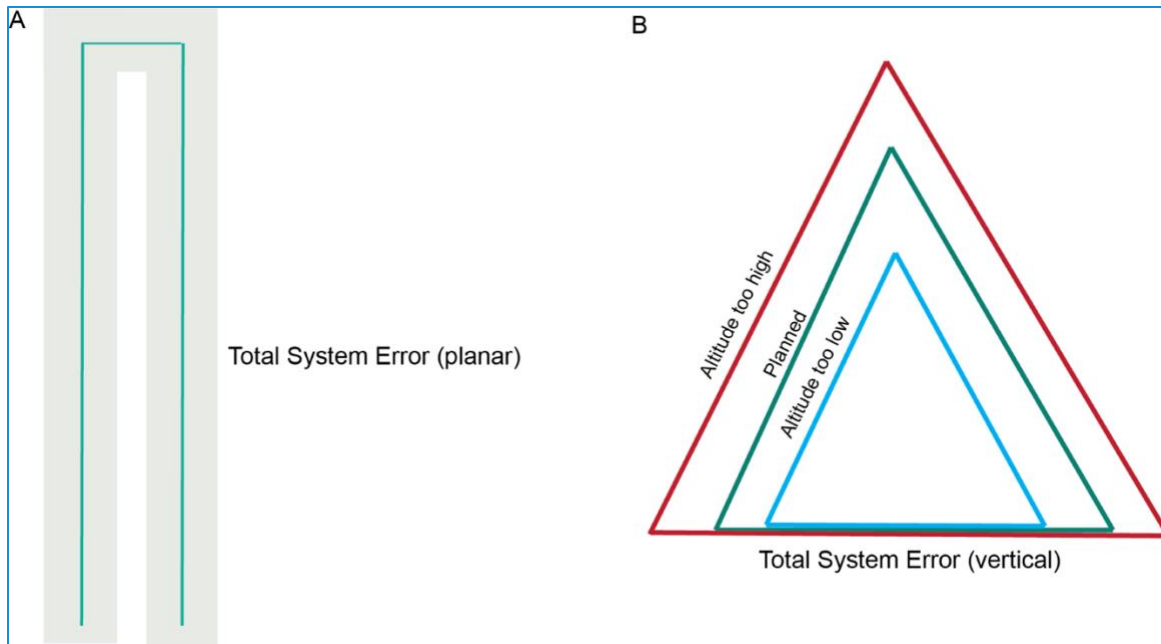


Figure 8. Example of planar (A) and vertical (B) Total System Error. In A, the green line represents the desired UAV flight path. The shaded area is the combination of the navigation solution error, flight technical error, and path definition error. This planar view illustrates how the position of the flight lines can be uncertain in comparison to the desired flight path based on the planar Total System Error. In B, the impact of the altimetry system error, path definition error, and flight technical error on the desired altitude (planned). The widths of the base of the red and blue triangles illustrate how the width of a flight line can be wider (red) or narrower (blue) in comparison to what was planned (green) depending on the vertical Total System Error. In addition, higher flight altitudes will result in larger pixels than planned while lower flight altitudes will result in smaller pixels than planned.

The best quality HSI results from the UAV and gimbal working together to minimise variability in attitude. The use of a gimbal reduces distortions in the raw imagery over data collected with hard mounted sensors (see Lucieer et al., 2014; Arroyo-Mora et al., 2023) and should be a standard implementation for UAV operations aiming to validate satellite products wherever possible. This is likely to be important for the validation of fine resolution satellite data (< 1 m) but reduces as the resolution increases. Similarly, it will be more useful over heterogeneous sites. In cases where the use of a gimbal is not possible, the overall stability of the UAV becomes even more important and efforts should be made to generate well geometrically corrected imagery, and at minimum report the IMU attitude data output (i.e. yaw, pitch, and roll). Ultimately, UAV stability and minimisation of attitude parameters such as roll influence the number of flight lines required. When using a stable system as shown in Figure 9B, less overlap between neighbouring flight lines (e.g. 20% without view zenith angle filtering, 55% with) can be planned without risk of gaps of no data between lines. In contrast, an unstable system or one not using a gimbal requires a larger overlap (> 35% without view zenith angle filtering, ~70% with) because roll experienced by the sensor can lead to imaging gaps between flight lines (Figure 9A). Over large areas, a larger overlap between lines can result in additional flight lines (and therefore time) being required to adequately cover the area.

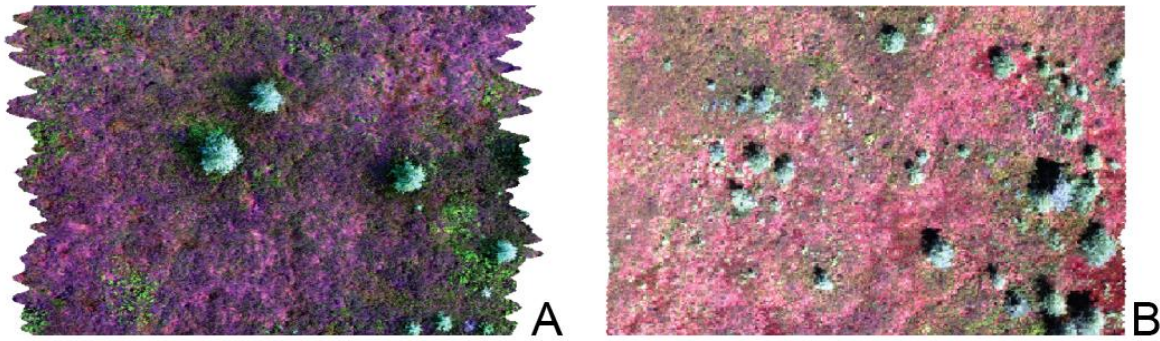


Figure 9. A) UAV HSI flight line with substantial influence of UAV roll as is often seen in imagers hard mounted to the airframe. The influence of roll remains in the image following geocorrection as can be seen by the uneven edges of the flight line. In such cases the overlap between lines needs to be increased (> 25%). B) UAV flight line following geocorrection where the HSI experienced minimal roll during the flight as indicated by the nearly straight edges of the image. The overlap between lines can be decreased to 10% in such cases.

6.4.3 Flight planning software

With the diversity of heavy lift UAVs and advanced hyperspectral sensors, flight planning tools vary among different platforms (e.g. DJI Pilot 2, UgCS, Mission Planner) and HSI sensor manufacturer recommendations (e.g. HySpex, Specim, Headwall, etc.). A large component of the total system error is the path definition error caused by the user during flight planning. The path definition error is the difference between the intended flight path (horizontally and vertically) and the programmed flight path (Transport Canada, 2021). Contributing factors can be the result of map projection differences, Earth reference model differences, altitude considerations, and inclusion of a terrain model. Choosing an appropriate flight planning tool can help reduce this component of the total error.

A baseline tool should provide the user an input with the required flight parameters, i.e. nominal altitude, speed, and required flight line overlap, in addition to sensor characteristics such as the FOV, number of across tracks pixels, and integration and frame times. More advanced tools allow the user to simulate the energy contribution to each pixel in the image and may also allow for planning based on specialised parameters such as a necessary signal-to-noise ratio (SNR) and vary the definition of a pixel (e.g. pixel spacing versus percent contribution of materials). Ideally, the output of the flight planning software generates key estimated parameters such as pixel resolution and spacing (x, y), flight line width, adjacent line separation, and overlap. An example of a flight UAS HSI planning tool (HyPlanT) developed by Naprstek and Inamdar (2022) is shown in Figure 10. The output from HyPlanT also generates a graph of the point spread function (PSF), that provides the user with an estimation of the energy distribution within the pixel (for more details see Inamdar et al., (2020)).

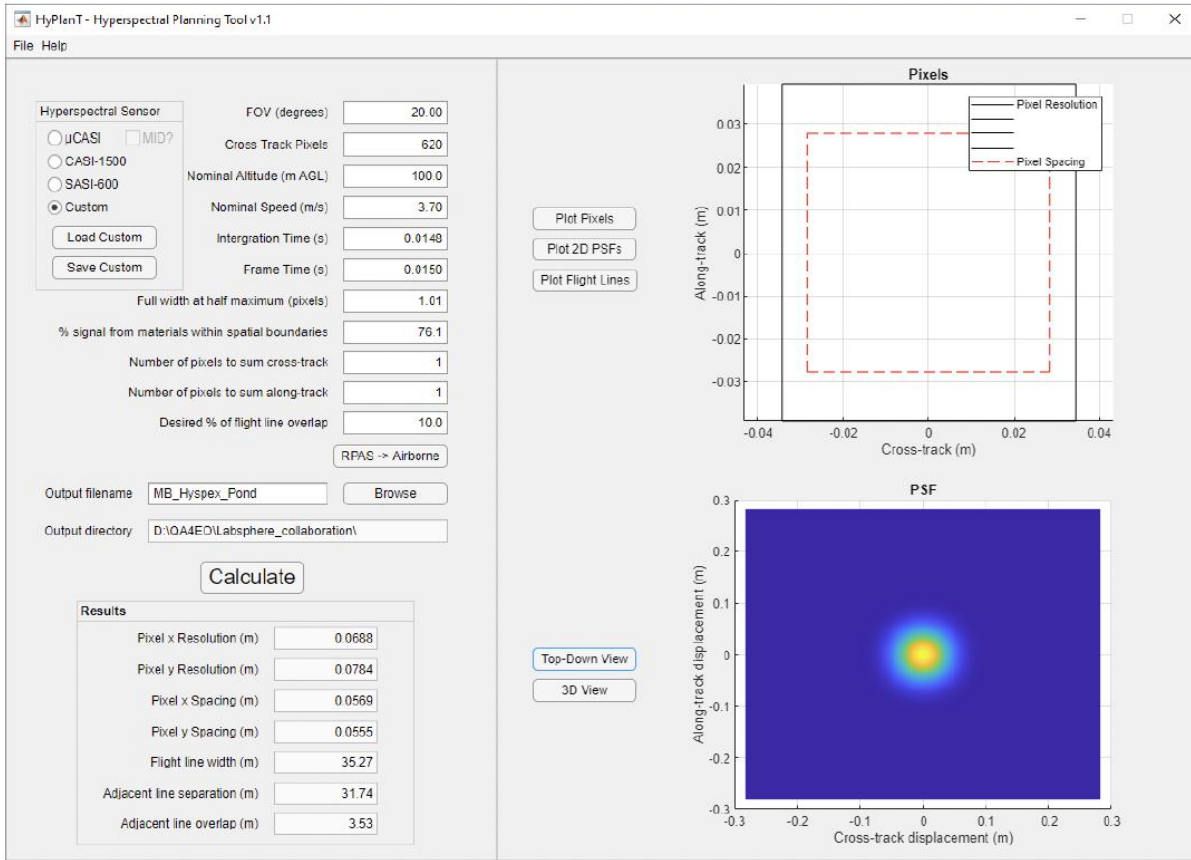


Figure 10. Example screen of HyPlanT flight planning tool for HSI data collection for UAS and airborne platforms.

Population of the flight parameters is critical in acquiring the correct data for the validation of satellite products. *It is recommended* to set the orientation of the flight lines to be aligned (parallel) with the overpass angle of the satellite being validated [R14]. This reduces the impact of surface anisotropy in the azimuth dimension by replicating the observation geometry of the satellite. Similarly, the overlap percentage (the proportion of the neighbouring flight line measured) is a key parameter because it facilitates the reduction in data gaps, but more critically because it can be used to increase the sampling of an area when conducting view zenith angle filtering.

Another aspect of the flight plan is the development of the spatial layers containing the actual flight lines which will be input into the flight control system. The most common formats are ESRI shapefiles and KML/KMZ. Developing the flight lines in Geographic Information System (GIS) software like QGIS or ArcGIS *is recommended* [R15], as these provide the user with the choice of precise coordinate systems (geographic or projected). Given the precision needed from UAS-based satellite validation, the use of Google Earth as a flight planning tool is inadequate. Online basemap services collect imagery from third parties and their imagery is refreshed unevenly across the globe, with remote, less popular areas updated less frequently resulting in imagery

that can be years out of date (Kalacska et al., 2020). Furthermore, Google Earth uses a Mercator projection based on a spherical datum, the WGS 1984 Web Mercator coordinate system (EPSG:3857). This is different from the World Geodetic System (WGS84) standard for global positioning systems relied upon by UAV flight controllers (EPSG:4326). Depending on the datum or projection in which incoming flight planning data are provided, the difference without a proper datum transformation or reprojection can result in errors up to several meters. Properly implemented, GIS software can minimise conversion errors resulting from differences in coordinate systems.

The coordinate system or projection information is also critical to correctly determine the position of a mobile base station for RTK corrections carried out by the flight controller. If the base station position is determined based on incoming Networked Transport of Radio Technical Commission for Maritime Services (RTCM) via Internet Protocol (NTRIP) corrections, a conversion may be necessary. For example, a caster (a server that provides the real-time correction (RTC) data to the NTRIP clients) in a European Union country may be using a local grid, whilst the flight controller is expecting the flight to be planned in WGS84. A conversion tool such as the [ETRF/ITRF Coordinate Transformation Tool \(ECTT\)](#) would be required in these instances. Several such coordinate conversion tools exist with precise conversions based on the reference system used in the region/country.

Another aspect to consider for planning precise flight lines is that geographic coordinate systems such as latitude/longitude with the WGS84 datum are angular measures and as such the distance on the ground for the same angular measure is not equal between the N-S direction and E-W direction, anywhere other than on the equator. For example, one degree of latitude is approximately 111 km, whereas one degree of longitude only equals 111 km at the equator; towards either pole the distance decreases. At 45° North, one degree of longitude equals ~78 km. Conversely, projected coordinates allow for the preservation of certain properties such as distance (equidistant), local angles (conformal) or area (equal area).

Figure 11 provides an example of a flight plan used for the validation of the German Environmental Mapping and Analysis Program (EnMAP hyperspectral satellite) overpass of the Mer Bleue Peatland Observatory Cal/Val Supersite. This flight plan was developed in ArcGIS Pro v. 10.7.1, in the Universal Transverse Mercator (UTM) projection (UTM 18N WGS84), generating ESRI shapefiles, which were reprojected to a Geographic Coordinate System with WGS84 datum and imported into the flight controller. Upon selecting the flight plan, the user can assess the total calculated flight time and estimated battery use (in the case of DJI drones) and develop logistics such as battery charging requirements. For instance, in the example in Figure 11, the heavy lift UAV carrying a HySpex VS-620 payload (~6.7 kg without gimbal) required ~9 minutes to collect imagery at 3.7m/s and 100 m above ground level (AGL) for each area.

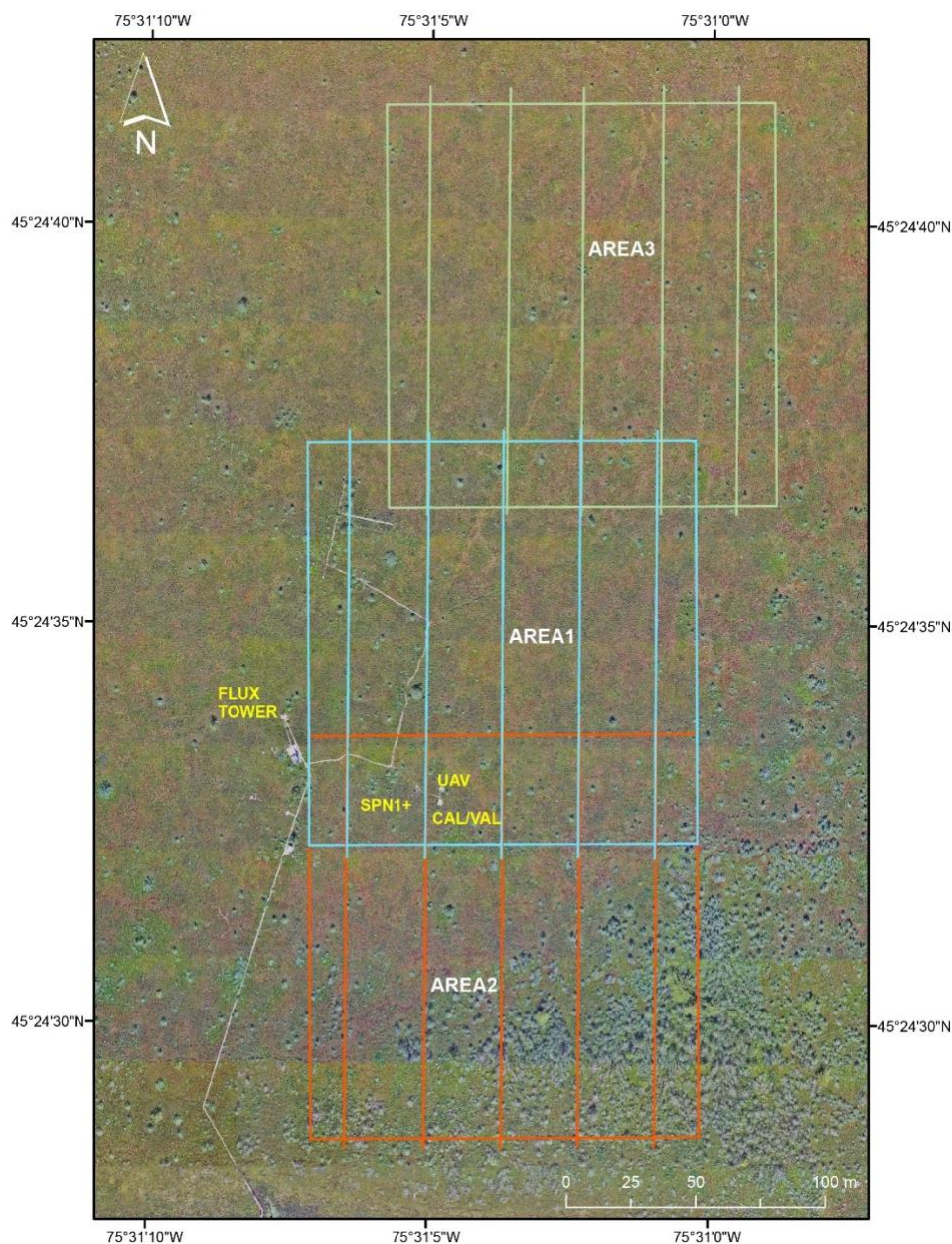


Figure 11. Example of a flight plan for UAV-HSI based EnMAP reflectance product validation at the Mer Bleue Peatland Observatory.

6.4.4 Advanced applications – estimation of spatial coverage

The final component related to flight planning addressed here is the assessment of the spatial coverage of the pixels within the image given specific acquisition and sensor characteristics. Use of such tools is two-fold, first tools such as HyPlanT can be used to assess changes in pixel PSF in response to changing flight or sensor characteristics. Figure 12 illustrates the impact of changing flight speed (with all other variables constant) on the PSF. As the flight speed increases, the pixel size in the along-track direction increases relative to the across track direction resulting in

elongated pixels. Keeping the speed constant and increasing the altitude results in changes in both shape of the PSF (elongated in the along-track direction to elongated in the across-track direction) and overall PSF area (larger at higher altitude).

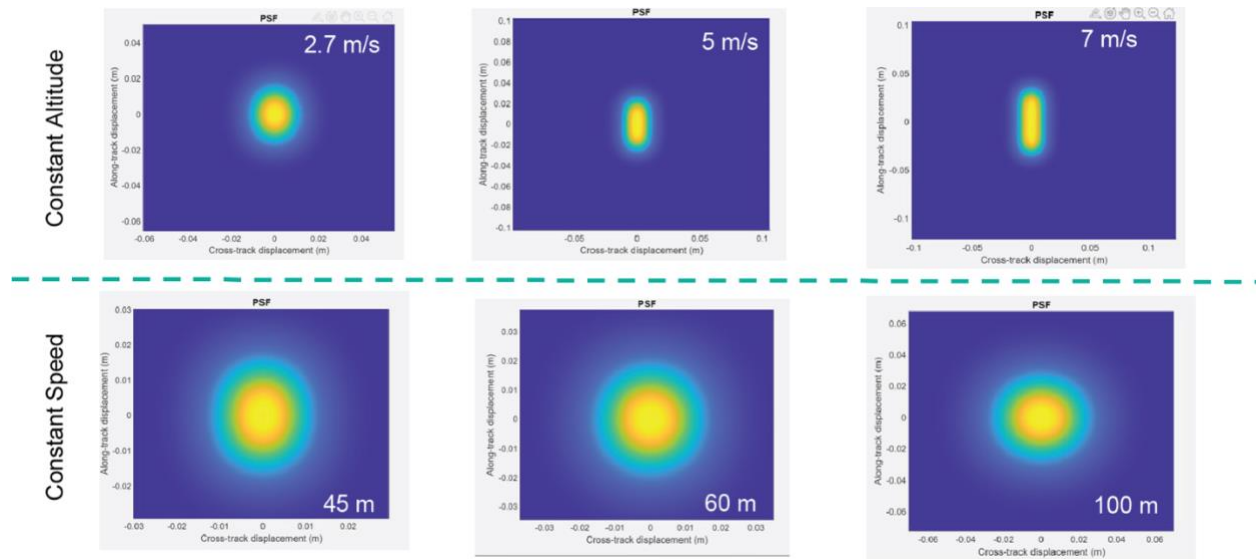


Figure 12. Impact on the PSF of increasing flight speed while keeping the altitude and other parameters constant (upper panels) and increasing altitude while keeping the flight speed and other parameters constant (bottom panels).

To fully appreciate the importance of the shape of the PSF, and the final pixel arrangement following geocorrection, an assessment of the spatial contribution of neighbouring pixels is warranted with a tool such as the *Spatial Coverage Map and Resampling Error Assessment* developed by Inamdar et al., (2023), which allows the user to estimate the percent of pixel loss, duplication, and spatial shift given a specific resampled pixel size following Inamdar et al., (2021). Pixel loss (PL) is defined as the total percentage of pixels from the original HSI dataset that were not used in the final resampled data product. Pixel duplication (PD) is defined as the total portion of pixels in the resampled raster (i.e. geocorrected image) that are duplicates of one another. Figure 13 illustrates the trade-off between PL and PD at different flight speeds and altitudes for a pushbroom HSI. In general, when PL is high, PD is low and vice versa. At all altitudes, as speed increases, PL decreases and PD increases. With the characteristics of the sensor used to generate Figure 13, the ideal acquisition scenario is 100 m AGL at a speed of 2.5 m/s, which preserves many of the acquired pixels, while minimising duplication. This results in a pixel loss of 6.3% and a pixel duplication of 9.6%.

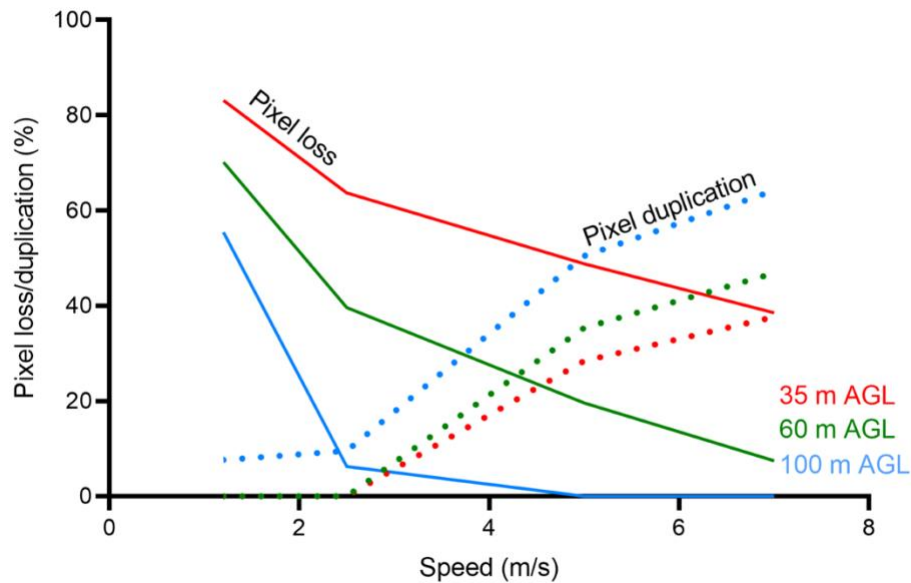


Figure 13. Relationship between flight speed and pixel loss and duplication at different flight altitudes for a given pushbroom HSI with specific characteristics (i.e. FOV, number of cross track pixels, integration time, frame time, FWHM, and spatial pixel summation).

Another factor to consider is the extent of the pixel being validated. Satellite pixel resolutions range from < 1 m to > 1 km. It is already clear from the previous sections that validation at the finer resolution can be done by UAV systems over many pixels in one acquisition whilst also maintaining square pixels and circular PSF shapes. This means that the characteristics of the site are less important if it is sampled fully (i.e. wall to wall). However, at the coarser scale this becomes much more challenging. Recommendations for expanding the extent of the potential satellite product resolutions are to select suitably homogeneous areas (or areas where the natural variability can be sampled) and using UAV-based hyperspectral systems to sample the area. This approach would accept that there will be data gaps, at the expense of covering the full area. To accomplish this, there may also be a need to fly higher and faster than is typically recommended. Nevertheless, the reflected radiance is still being captured from a portion of the surface located within the satellite pixel, despite the UAS pixel shape and projection on the surface not conforming to that needed to produce imagery free of visual artefacts. For all resolutions being targeted, the definition of the ROI should include a spatial buffer to account for heterogeneity of the surroundings and adjacency effects (Ong et al., 2018).

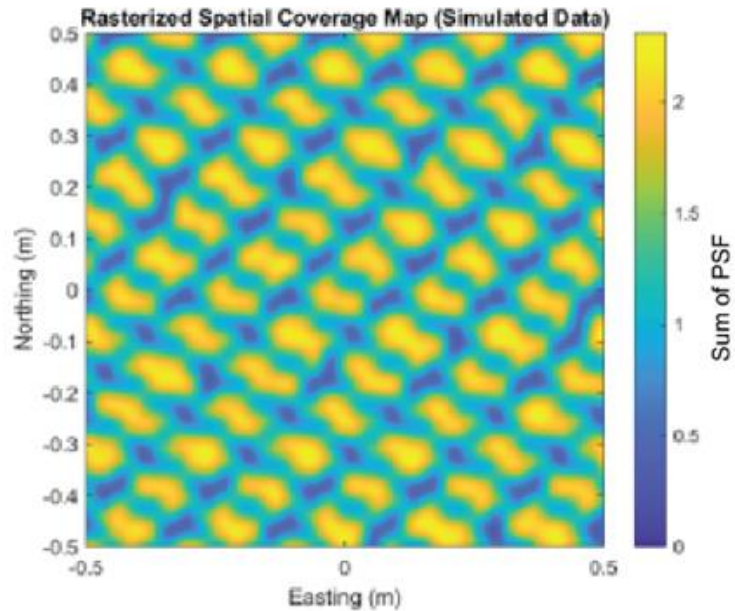
Returning to one of the initial statements in this section, “ultimately, flight planning is a needs-driven approach where one determines the required characteristics of the imagery, then works backwards based on the sensor and flight geography to determine the appropriate combination to achieve those characteristics”, Figure 14 illustrates the difference in the data product (i.e. geocorrected raster) if the same flight acquisition parameters are used with two different

hyperspectral imagers that have different characteristics; this emphasises that there is no single set of flight parameters that applies to all imagers.

Sensor 1

FOV: 20°
Integration time: 5.6 ms
Frame time: 5.83 ms
XTrack pixels: 640
Altitude: 60 m AGL
Speed: 5 m/s

Resolution X: 4.0 cm
Resolution Y: 4.4 cm



Sensor 2

FOV: 34.2°
Integration time: 9.0 ms
Frame time: 11.0 ms
XTrack pixels: 1833
Altitude: 60 m AGL
Speed: 5 m/s

Resolution X: 2.4 cm
Resolution Y: 4.0 cm

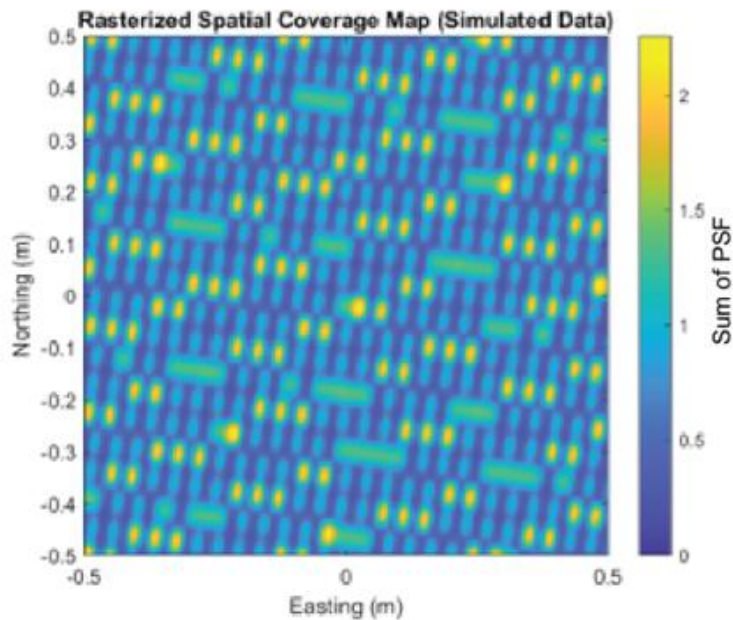


Figure 14. Spatial map of the sum of the PSFs of the pixels within an HSI dataset quantifying how HSI data spatially samples spectral information across an imaged scene (Inamdar et al., 2023).

6.5 Reference targets

Reference targets are commonly deployed as a means of assigning known reflectance values for a scene (i.e. calibration) or for checking the calibration of reflectance values acquired from UAV-based sensors (i.e. validation). Ideally, candidate targets would be spectrally smooth, spatially uniform, Lambertian targets of adequate size to cover many image pixels in the along/across-track directions. As materials with perfect Lambertian properties are not possible, a few possibilities exist for optimal deployment: 1) the imager views the reference targets over the full range of observation geometries that will be used in the final image, or 2) the anisotropic properties of the reference target should be measured. Since the maximum time of a UAV flight is typically between 10 min – 25 min, it may be possible to assume that the change in reflectance resulting from the changing illumination geometry is sufficiently small that it can be considered negligible – however, this will be location (with sites closer to the equator experiencing larger solar geometry changes in a shorter period of time) and target dependent (with some targets being closer to Lambertian than others).

6.5.1 Panel substitution method (PSM)

A technique called the Panel Substitution Method (PSM) is commonly implemented for calibrating the pixels of the UAV-based sensor; it requires the reflectance of the reference target to be known. The ratio of the signal observed over the target of interest and that observed over the reference panel (both in digital numbers) is scaled by the calibrated spectral reflectance of the reference – equation 2.

$$R_{HCRF,t}(\theta_s, \varphi_s, 2\pi; \theta_v, \varphi_v, \omega_v; \lambda) = \frac{S_t(\theta_s, \varphi_s, 2\pi; \theta_v, \varphi_v, \omega_v; \lambda)}{S_r(\theta_s, \varphi_s, 2\pi; \theta_v, \varphi_v, \omega_v; \lambda)} R_{HCRF,r}(\theta_s, \varphi_s, 2\pi; \theta_v, \varphi_v, \omega_v; \lambda) \quad 2$$

Where $R_{HCRF,t}$ is the hemispherical-conical reflectance factor (HCRF) of the field target, $R_{HCRF,r}$ gives the HCRF of the reference target, S_t gives the dark corrected field spectrometer radiance of the field target material (in digital numbers), S_r gives the dark corrected field spectrometer signal of the field reference material, θ_s is the solar zenith angle (SZA), φ_s is the solar azimuth angle (SAA), θ_v is the view zenith angle (VZA), φ_v is the view azimuth angle (VAA), ω_v is the solid angle defining the viewing optic AFOV, and λ is the wavelength. The following assumptions are made when applying this equation: the spectrometer is temporally stable and radiometrically linear (or has been corrected for non-linearity) with any stray-light correction applied; the signal is corrected for offset signal, most notably the dark current; the illumination conditions remain stable between the acquisition of the field reference and the field target data; and the reference target is Lambertian if observations of the field site will be made in angles other than those in place when viewing the reference target.

The PSM has a complexity versus uncertainty trade-off with lower uncertainties only achievable through additional measurements and calibrations. The benefits and problems with the implementation of each of the methods, in order of complexity, are identified below.

6.5.1.1 $R_{HCRF,r} = 1$

Setting the HCRF calibration value to unity is equivalent to assuming that the target is perfectly Lambertian across the full spectral range. This can provide results to be used to monitor relative changes between various targets acquired under similar illumination conditions. In general, this technique is not recommended, particularly for Cal/Val activities, and effort should be expended to ascertain the spectrally dependent calibrated reflectance, and to accommodate the non-Lambertian nature of the reference panel.

6.5.1.2 *Manufacturer calibration*

The reflectance quantities for manufacturer calibrations are not the same as the reflectance experienced in the field since the illumination conditions and mode of measurement will be different. Whilst it can be used as a first order proxy, this can lead to significant biases for some illumination angle conditions and is not recommended. The calibrated quantity is spectrally dependent. Application of this parameter as $R_{HCRF,r}$ in equation 2 will therefore provide results with appropriate spectral features. It is worth noting that both the absolute levels and spectral shape can change with use of the reference panel over time. Using this approach assumes that the reference target is Lambertian.

6.5.1.3 *SZA compensation*

Regardless of the overall approach to the determination of the target HCRF, a first order correction to compensate for the systematic change in the downwelling irradiance levels due to changes in the SZA can be applied. This correction can have a positive impact on the quality of the $R_{HCRF,t}$ results should the atmospheric conditions be stable but may be less helpful, or even problematic, if non-systematic changes such as those encountered due to dynamic cloud or haze conditions are more significant than the systematic changes. The magnitude of the correction will be impacted by the length of time between the reference and target scans. The sensitivity to this time delay will be dependent on the time of day that the measurements were acquired. The change in SZA as a function of time will be negligible at solar noon, with the rate of change increasing the farther the time of acquisition is from solar noon.

Two approaches may be applied to perform a first order correction addressing the systematic changes in the downwelling irradiance conditions between the reference and target scans:

- 1) A time-based interpolation of a pair of reference scans, acquired immediately prior to and following a sequence of target scans, can be used to generate a temporally-adjusted estimate of S_r in equation 2 (this is used in Origo et al., 2020) for each individual target scan.

- 2) A cosine correction is capable of modelling changes in the downwelling irradiance levels due to the advancing position of the SZA between the reference and target field spectrometer scans. This is accomplished by adapting equation 2 to:

$$R_{HCRF,t}(\theta_s, \varphi_s, 2\pi; \theta_v, \varphi_v, \omega_v; \lambda) \quad 3$$

$$= \frac{S_t(\theta_v, \varphi_v, \lambda) \frac{\cos \theta_{s,r}}{\cos \theta_{s,t}}}{S_r(\theta_v, \varphi_v, \lambda)} R_{HCRF,r}(\theta_s, \varphi_s, 2\pi; \theta_v, \varphi_v, \omega_v; \lambda)$$

Where $\theta_{s,r}$ and $\theta_{s,t}$ are the solar zenith angles at the time of reference and target measurements respectively.

6.5.1.4 Calibration of reference target as BCRF

The dominant source of illumination over much of the solar spectral range, particularly on a clear sky day, is the direct component. The term $R_{BCRF,t}(\theta_s, \varphi_s, \omega_s; 0^\circ, \varphi_v, \omega_v; \lambda)$ therefore provides a suitable parameter to scale the signal ratio to $R_{HCRF,t}$, where ω_s is the solid angle of the cone defining the illumination. There are multiple approaches to this shown to be applicable to sintered polytetrafluoroethylene (PTFE), otherwise referred to as 99% Spectralon:

- 1) A measurement of $R_{BCRF,r}(0^\circ, \varphi_s, \omega_s; 45^\circ, \varphi_v, \omega_v; \lambda)$ combined with a spectrally independent normalised BCRF function which defines the relationship between the BCRF at a given SZA and that at an illumination angle of 45° . At least 3 independent studies were shown to provide a consistent nBCRF for different 99% Spectralon panel (Soffer et al., 2019). Note that the panels used in these studies were all relatively new samples and the impact of degradation on this assumption has not been assessed.
- 2) Should measurements of $R_{BCRF,r}(0^\circ, \varphi_s, \omega_s; 45^\circ, \varphi_v, \omega_v; \lambda)$ not be available, the results of a study of the reflectance properties of 99% Spectralon, published by NIST (Cooksey et al., 2015) can be applied to $R_{HCRF,r}(\theta_s, \varphi_s, 2\pi; 8^\circ, \varphi_v, \omega_v; \lambda)$ provided calibration data to estimate the $R_{BCRF,r}(0^\circ, \varphi_s, \omega_s; 45^\circ, \varphi_v, \omega_v; \lambda)$. A linear wavelength dependent relationship was identified between the two geometries between 400 nm and 2500 nm with a slope of $(-9.4 \pm 2.7) \times 10^{-6}$ and a y-intercept of 1.0222 ± 0.0039 . Note that this relationship is for 6°:h as opposed to 8°:h which is the typical geometry for which hemispherical reflectance calibration data is provided. However, the difference between two geometries is expected to be small as the view geometry approached 0° .
- 3) $R_{BCRF,r}(0^\circ, \varphi_s, \omega_s; 45^\circ, \varphi_v, \omega_v; \lambda)$ can be derived from multi-angular measurements of the BCRF. The reference panel would be characterised in the laboratory beforehand, and the information used to create the spectrally dependent nBCRF function discussed in point 1. This approach was used in Origo et al., (2020).

6.5.1.5 Direct-to-total weighting

As previously noted, $R_{BCRF,t}(\theta_s, \varphi_s, \omega_s; 0^\circ, \varphi_v, \omega_v; \lambda)$ can provide a suitable reflectance factor to describe situations where the direct component of the downwelling irradiance is dominant. For clear skies, the diffuse component can grow to greater than 50% for wavelengths below 400 nm.

To accommodate the diffuse component, a weighted estimation of $R_{HCRF,r}(\theta_s, \varphi_s, 2\pi; \theta_v, \varphi_v, \omega_v; \lambda)$ is recommended for the ambient illumination conditions (α):

$$R_{HCRF,t}(\alpha; \lambda) = (1 - d(\lambda)) \left(\frac{S_t(\theta_v, \varphi_v, \lambda)}{S_r(\theta_v, \varphi_v, \lambda)} R_{BCRF,r}(\theta_s, \varphi_s, \omega_s; \theta_v, \varphi_v, \omega_v; \lambda) \right) + d(\lambda) \left(\frac{S_t(\theta_v, \varphi_v, \lambda)}{S_r(\theta_v, \varphi_v, \lambda)} R_{BHR,r}(2\pi; 2\pi; \lambda) \right) \quad 4$$

Where $d(\lambda)$ is the diffuse-to-total irradiance ratio. If aerosol optical depth and water vapour information are available, $d(\lambda)$ can be computed using an atmospheric correction model (e.g. Vermote et al., 1997); this assumes a clear sky is present. This approach can be supplemented by using a broadband diffuse/direct irradiance instrument to scale the spectral diffuse-to-total estimates (e.g. as in Origo et al., (2020)). The ideal case is spectrally resolved estimates of the diffuse-to-total irradiance. Where field spectrometer data of the diffuse illumination conditions for a reference target have been collected, $d(\lambda)$ can be determined as (Schiller & Silney 2010):

$$d(\lambda) = \frac{S_{r,shaded}(\lambda) + (S_{r,unshaded}(\lambda) - S_{r,side-shaded}(\lambda))}{S_{r,unshaded}(\lambda)} \quad 5$$

Where $S_{r,unshaded}$ is the dark corrected field spectrometer data of unshaded reference panel in DN, $S_{r,shaded}$ is the dark corrected field spectrometer data of shaded reference panel in DN, and $S_{r,side-shaded}$ is the dark corrected field spectrometer data of side-shaded (shadowed area to the side) reference panel in DN.

$R_{BCRF,r}$ is derived as explained in the sections above, whilst $R_{BHR,r}$ can be derived in three manners:

- a. If comprehensive goniometric $R_{BCRF,t}(\theta_s, \varphi_s, \omega_s; \theta_v, \varphi_v, \omega_v; \lambda)$ measurements are available of the reference panel, $R_{BHR,r}(2\pi; 2\pi; \lambda)$ can be computed as from:

$$R_{BHR,r}(2\pi; 2\pi; \lambda) = \frac{\int_0^{\frac{\pi}{2}} \int_0^{2\pi} R_{BCRF,r}(\mathbf{0}^\circ, \varphi_s, \omega_s; \theta_v, \varphi_v, \omega_v; \lambda) \sin \theta_v \cos \theta_v d\theta_v d\varphi_v}{\int_0^{\frac{\pi}{2}} \int_0^{2\pi} \sin \theta_v \cos \theta_v d\theta_v d\varphi_v} \quad 6$$

- b. Additional bi-hemispherical calibration of the reference panel.
- c. If goniometric measurements or bi-hemispherical reflectance calibration are not available, $R_{HCRF}(\theta_s, \varphi_s, 2\pi; 8^\circ, \varphi_v, \omega_v; \lambda)$ can be used as a proxy.

Note that both these derivations assume isotropic diffuse downwelling irradiance where in fact the diffuse sky irradiance will be anisotropic.

6.5.1.6 Full sky weighting

The previous section notes that isotropic diffuse illumination is assumed as part of the formulation. A further improvement would be to characterise the sky radiance distribution such that it is known for the full hemisphere (ideally spectrally). This does not have to be an absolute

measurement. Combined with a characterisation of the reference target multi-angular reflectance factor (BCRF), this would facilitate the creation of a calibration coefficient for the reference target which corresponds with the exact illumination conditions at the time of the measurement.

6.5.2 Reference target characteristics

For the reference targets, a minimum of two radiometrically distinct (different reflectance levels) targets *are required* [R16]; one should be at or below the lowest HCRF expected in the field and one should be at or above the brightest. This ensures that the response of the hyperspectral sensor under evaluation is within the calibration range and that the linear relationship passes through the origin. Note that the lower reflectance target shouldn't be much lower than the minimum expected scene reflectance in order ensure noise does not dominate the measurement. Similarly, the bright target reflectance should be such that it avoids saturation. An assessment of the linearity should be carried out in laboratory conditions. Additional targets may be considered to further assess the magnitude of potential non-linearities in the response. This may be of particular benefit if extremely low-level targets (e.g. water) are a target of interest. The lower HCRF reference target should not be smaller than noise floor of the instrument. Bi-hemispherical reflectances of suitable reference targets would be around 5% – 10%. The greater HCRF reference target should not saturate the hyperspectral imager. Saturation levels will be dictated by numerous factors including downwelling irradiance conditions, sensor radiometric response, full well capacities, and system integration times. In general, it is possible to avoid saturation by decreasing the integration time. However, this comes at the expense of a reduced signal-to-noise ratio.

Reference targets should be as close to Lambertian and spectrally flat as possible to minimise sensitivity to illumination conditions. They should be large enough relative to the imager resolution at the height of the UAV platform so multiple “pure” pixels can be retrieved (i.e. free from adjacency effects and mixed pixels) in the along- and across-track directions. A minimum central area of 9 x 9 pixels from which to extract mean and standard deviation with a minimum perimeter of 3 pixels, to eliminate any edge effects due to spectrometer point spread function and scattered light issues, *is recommended* [R17]. For line imager UAV-sensors, the reference target should be larger to accommodate for the possibility of irregular sensor motion (roll, pitch, and yaw) during the acquisition of the reference target.

Cloth or tarpaulin materials are often considered as reference targets due to their ease of transport and deployment. Such materials are likely to have greater anisotropic properties when compared to hard targets due to the weave of the fabric and creasing caused during storage. Replicating the UAV-imager geometries with the ground instrumentation is particularly important in this scenario. In addition, every effort should be made to eliminate any slope or creases in the deployed fabric.

It is much easier to calibrate smaller reference targets in the laboratory compared to the relatively large targets required for UAV flights (i.e. > 0.5 m). As such, it is common practice for the calibration of a reference target to be transferred to a larger reference target in the field by using a field spectrometer. However, this requires the spectrometer being used to be temporally stable and radiometrically linear. It is recommended that the spectrometer has a narrow AFOV foreoptic to ensure that the retrieved reflectance factor is as close to the UAV quantity as possible. It is noted that in general a hyperspectral imager will have pixel AFOVs of $\ll 1^\circ$, while field spectrometer foreoptics will usually be $> 1^\circ$. Similarly, stable irradiance conditions are required for the duration of the flight. The reference target must be levelled as precisely as possible, and it must be ensured that reference targets are placed in locations where they can't be shadowed by objects in the scene and ideally there shouldn't be any tall objects in the scene that could spectrally contaminate the reference target spectra (e.g. tall trees nearby blocking skylight to the reference target). This also applies to the operator who should use an extended arm to conduct the measurement as far away from their body as possible, thereby minimising their influence on measurement. They should also primarily wear dark clothing and avoid shiny materials (Goetz, 2012). However, it is noted that there will always be some portion of the sky light that will be blocked by the operator, as such this should be consistent for all the measurements by ensuring the operator is blocking the same portion of sky light per measurement.

Monitoring of the calibrated reference panel over the course of the measurement campaign is desired since the panel can become contaminated with dirt/dust or from ultraviolet radiation degradation. The procedure to do this involves conducting measurements at the start and end of each day with an equivalent panel that is only exposed for these measurements. A spectrometer can be used to conduct near-simultaneous measurements of both panels in this scenario. It is important that both panels are at the same height (and levelled) and both measurements are taken from the same height (and levelled) to ensure a consistent measurement configuration. Nevertheless, it is advantageous to only use the calibrated reference panel when conducting measurements. Ensuring it is covered when not in use will minimise contamination.

6.6 Post-processing

The planning and implementation of the UAV-based hyperspectral flights is half of the work required to produce surface reflectance data consistent with the satellite product. The other half is processing that data from the raw information provided by the imaging system and supporting instruments into the quantity that can be readily compared with the satellite product. This section lays out the necessary steps to produce this information.

Where dark subtraction does not take place automatically within the processing, this should be applied. For example, some systems require that users make dark measurements prior to the

UAS flight. Where calibration and characterisation information (e.g. uniformity correction, bad pixel mapping, etc.) is present, this should be applied to the dark subtracted data.

6.6.1 VZA filtering

View zenith angle filtering is the process of removing data from the retrieved hyperspectral data cube that corresponds with parts of the instrument’s AFOV which are significantly different from that observed by the satellite overpassing the field site. This can be done for pushbroom and 2D imagers by assigning the along-track pixels view zenith angles based on the overall AFOV of the foreoptic divided by the number of along-track pixels. For example, Origo et al., (in preparation) applied a view zenith angle filter of $\pm 2.5^\circ$ around the $\sim 6^\circ$ view zenith angle at which Sentinel-2 observes the Barrax field site. This same principle was applied as a condition of the defined protocol experiment in SRIX4Veg. The benefit of filtering by view angle is to remove the contribution of reflectance from view angles not viewed by the satellite sensor. Over targets with strong anisotropic properties (such as vegetation), the reflectance from out-of-AFOV reflectance can dramatically increase or decrease the overall reflectance. It is *recommended* that this be applied even if it results in data gaps (assuming the target area is reasonably homogeneous) [R18]. It should be noted that the flight planning should typically consider whether this type of filtering is likely to take place, and often account for it by increasing the overlap beyond the values typical of providing wall-to-wall pixel coverage (10% gimbaled/25% hard mounted). The exact increase depends on the filter size with finer angular bins requiring greater overlaps. Figure 15 demonstrates the importance of correcting for BRDF effects or observing in the same geometry of the satellite sensor at an “ideal” field site (i.e. selected due to low anisotropy and spatially homogeneous).

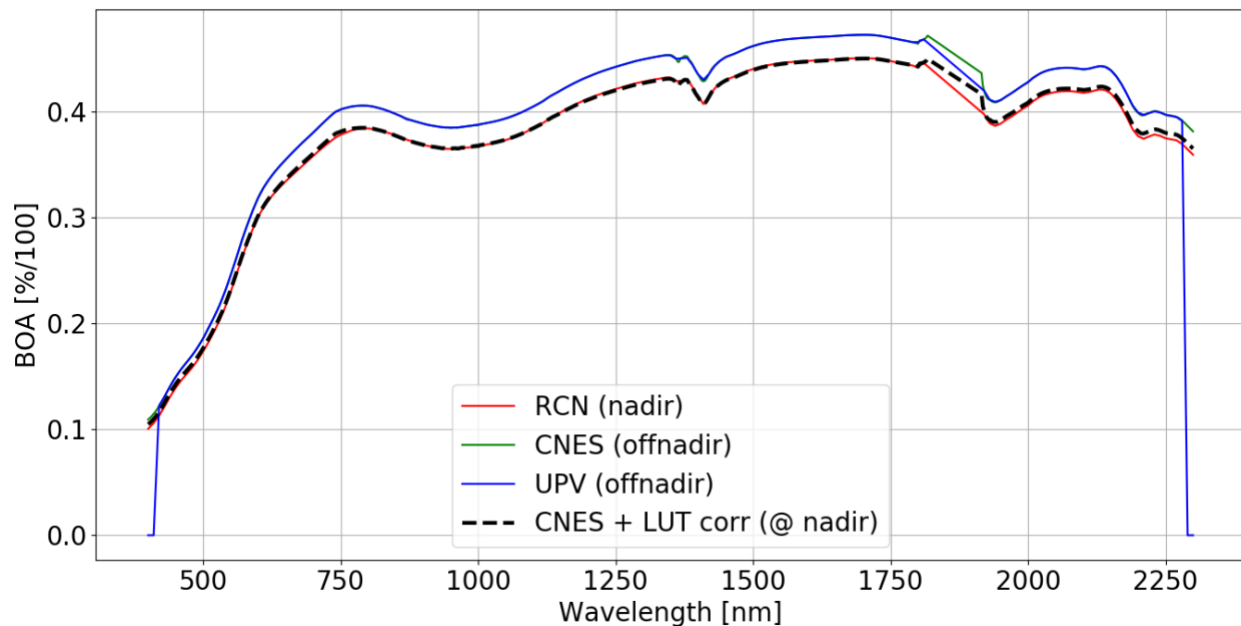


Figure 15. BOA surface reflectance at the Gobabeb RadCalNet site (Bouvet et al., 2019) at nadir (red), at simulated off-nadir (green), adding BRDF LUT bias (blue) to the nadir observations and correcting the off-nadir simulations with the BRDF bias (dashed

black line). For visibility reasons, the uncertainty of ~5% has been excluded from all the surface reflectance spectra (Pflug et al., 2023).

6.6.2 Hyperspectral data aggregation

Based on the points mentioned in previous sections around pixel loss and duplication, *it is recommended* that hyperspectral data is orthorectified but not converted into a raster format (i.e. gridding or stitching) [R19]. This preserves all pixels measured in the field (which improves estimates of the mean and standard deviation of the reflectance in a particular area of the image due to increased sample numbers) and provides memory and performance advantages (Inamdar et al., 2021). In addition, the duplication of pixels (e.g. using nearest neighbour) or combinations of multiple pixels (e.g. bilinear, bicubic, etc.) creates correlation within localised parts of the image which then needs to be tracked and accounted for in the uncertainty calculations (as well as reducing the number of independent measurements). This is particularly the case for pushbroom sensors, however, even 2D imagers will experience this effect when the images are stitched together. Point spectrometers will already create data in a point cloud format.

The hyperspectral point cloud should be converted to band-integrated values before conversion to reflectance (Burggraaff 2020). Here the spectrally resolved data is convolved with the spectral response function of the band of the satellite sensor of interest. This is expressed in equation 7:

$$S(\lambda_b) = \frac{\int_{\lambda_{b,min}}^{\lambda_{b,max}} S(\lambda) T_b(\lambda) d\lambda}{\int_{\lambda_{b,min}}^{\lambda_{b,max}} T_b(\lambda) d\lambda} \quad 7$$

Where $S(\lambda_b)$ is the band integrated signal, $S(\lambda)$ is the spectrally resolved signal at wavelength λ , $T_b(\lambda)$ is the spectral response of the satellite sensor band, and $\lambda_{b,max}$ and $\lambda_{b,min}$ are the maximum and minimum wavelengths in which that band is sensitive.

Once the hyperspectral point cloud is converted into the geospatial coordinate system of the satellite product it should be filtered based on the coordinates of the boundary of the pixel being validated. This ensures that every pixel (passing through the previous filters) falling in this location contributes to the UAV-based estimate of the surface reflectance of the satellite pixel. The spatially-averaged band reflectance for the pixel should be the final outcome.

6.7 Calibration and characterisation

The importance of instrument calibration and characterisation in the role of traceable in situ measurements is clear (Goryl et al., 2023), with the third defining characteristic of a FRM being:

“A comprehensive uncertainty budget for all FRM instruments and for the derived (processed) measurements (including any transformation of the measurand to match the measurand of the satellite product) is available and maintained.”

In the case of this protocol, the instrumentation in question is the hyperspectral camera, whether this is a point-imaging system, snapshot or a pushbroom system, and all supporting instrumentation such as the field spectroradiometers and reference targets. Calibration and characterisation are important factors in developing an understanding of the measurement system in question and unlike lab-based measurements, in situ field measurements often subject the instrumentation to additional uncontrolled environmental influences such as temperature extremes (heat and cold), humidity, and turbulence (motion during travel and setup). With UAV-based instrumentation, weather conditions and pilot experience can lead to heavy landings which can lead to changes in the system response. Similarly, shorter scale system responses related to temperature changes are also likely, especially where cooling of the detectors is not available. It is these changes that are particularly important to understand and record, with regular calibration and characterisation, between campaign seasons, *required* to fully understand the uncertainties pertaining to all instruments and artefacts [R20].

The three most used types of hyperspectral UAV instrumentation are point-imagers, 2D frame imagers and pushbroom imagers, each of which can have varying spectral range and resolution. Point imagers generally have high spectral resolution and low weight benefits but lack spatial referencing and do not provide spatially resolved data. In contrast, pushbroom imagers rely on the spatial referencing of consecutive image scanlines to form a cube of hyperspectral information which is highly dependent on the accuracy of the positioning and orientation (attitude) information. Methods to assess performance variations of the system using in-scene approaches is of critical importance.

For all types of sensors, spectral characterisation is important, with a need to understand and check the wavelength scale and spectral resolution to ensure the spectral response of the detectors are as they should be. This can be performed using spectral line emission lamps which emit narrow emission lines at known wavelengths against which the instrument spectral response can be assessed. These spectral peaks are plotted and compared to a modelled fit (traditionally Gaussian), from which the spectral position and Full Width at Half Maximum (FWHM) of this peak are determined; any offset from the expected value is noted and can be corrected for. Monochromators may also be used to produce monochromatic light that illuminates the detector, and by comparison with a reference detector, from which the spectral response function can be retrieved. In some scenarios the manufacturers will provide calibration and characterisation information which can provide a starting point for which future changes can be referenced to. In all cases it is important to ensure that the illumination of the spectrometer FOV is the same as when it is used in the field.

In addition to the spectral characterisation information mentioned above, certain additional measurements are fundamental in understanding the performance of a spectral imaging system so that the resulting data can be properly processed. This includes characterising the uniformity

in the response of the pixels across the image, the radiometric linearity, and identification of any bright/dead pixels. Ideally, these are measured using an absolute radiometric source such as an integrating sphere, where the full AFOV of the detector can be filled by a uniform light-source of known output. Following intensive use of a system, the level and distribution of pixel response would also be expected to change as the system ages and would require correction and monitoring over time, often in the form of pixel mapping. Similarly, characterising the impact of stray light of the imaging system is fundamental. Understanding of the polarisation sensitivity of the instrument may be required depending on the nature of the surface being investigated with highly specular surfaces like water likely to produce non-negligible errors.

For systems that rely on spatial referencing, the spectral and spatial behaviour of the pixel position and spatial lines is also an important consideration. Where there is a shift in the central wavelength as a function of the spatial position of the pixel within the focal plane, this is known as spectral co-registration (commonly referred to as ‘spectral smile’). Related to this is the spatial co-registration (commonly referred to as ‘keystone’), where instead there is a bending of the spatial response as a function of its position in the spectral axis (Figure 16). With the complexities of pushbroom systems, understanding these effects and being able to correct for them is important.

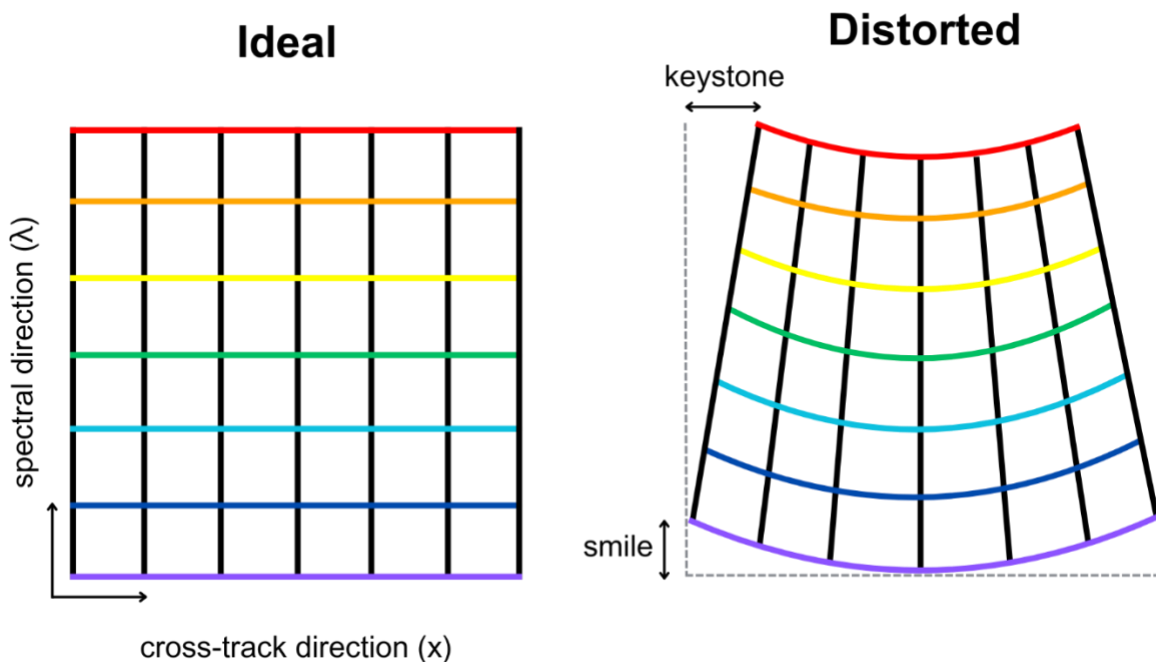


Figure 16. The ideal geometry of the spatio-spectral detector in a pushbroom imaging system (left). The impact of shifts in the spectral and spatial coregistration (right). This figure is reproduced from Yokoya et al., (2010).

For both the panel-based and irradiance-based in-situ methods of deriving reflectance, it is also important to consider the instrument response to dark measurements to eliminate the

background noise of the system. Many systems will include an automatic dark collection immediately prior to, or following, the exposed target measurements. If this is not present, *it is required* that manual dark readings should be performed before and after the flight [R21]. In such a situation, care must be taken to ensure that dark current measurements are performed at sufficiently frequent intervals to ensure optimal dark current correction of the target scans.

For systems that use the stop-point-shoot method of measurement, the positioning accuracy is an important source of uncertainty, the significance of which depends on the angle of measurement. For pushbroom systems there are uncertainty contributions from the continuous forward motion of the UAV arising from instabilities in the sensor orientation that occurs during the image acquisition. Attitude parameters (roll, pitch, and yaw) must be considered and corrected during orthorectification.

The calibration requirements depend on the method applied to derive surface reflectance. For the reference target-based approach, the traceability is usually via the calibration of the reference panel (as in Origo et al., 2020). This means that the spectrometer used to produce the reference panel HCRF values in the field does not need an absolute calibration of the detector response (Hueni et al., 2017). The only requirement is that it is radiometrically stable, responds linearly across different measurements and has a stable calibrated wavelength scale. For example, changes in temperature between successive measurements are kept to a minimum (e.g. through detector cooling if available or monitored if not) and the detector response to the observed spectral radiance level is linear (i.e. radiometric linearity). For irradiance-based surface reflectance measurements, an absolute calibration is required on both the radiance and irradiance sensors as the traceability source is through these.

Having considered the uncertainty contributions from the pre-campaign activities, it is then important to consider the contribution from measurement procedures in the field and the post-processing activities that follow, to cover the FRM requirement for traceability and full uncertainty budget (Goryl et al., 2023). The section below on uncertainty characterisation and propagation discusses this in more detail.

6.8 Uncertainty characterisation and propagation

Uncertainty analysis is the review of all sources of uncertainty and the propagation of that uncertainty through the measurement model. It is based on recommendations in the Guide to the Expression of Uncertainty in Measurement (GUM) (JCGM 2008). The GUM was developed by the JCGM (Joint Committee for Guides in Metrology), a joint committee of all the relevant standards organisations (e.g. ISO) and the BIPM (Bureau International des Poids et Mesures) and considers uncertainty as a distribution around the known measured value of quantity values that could be attributed to the measurand (the quantity intended to be measured). The intention is that the true, unknowable value lies within the stated uncertainty. *It is required* that the

uncertainty of all processed data is provided with the results and that the calculation of that uncertainty follows the guidance in this section [R22].

Key to deriving uncertainty is the measurement model which enables an estimate of the measurand to be provided and an associated standard uncertainty evaluated. The measurement model should include all quantities that affect the measurement result – this includes all post-processing and combinations with other uncertainty sources. The section below outlines the recommended method of undertaking this analysis and some of the considerations that can be applied for UAV-based field measurements.

Understanding the instrument’s unique characteristics and limitations is important in defining the overall uncertainty of the measurement. However, to understand the uncertainty of the final output each contributing factor must be included to develop the uncertainty propagation equation. As mentioned in the previous section, although the characteristics of the measurement instrument are important, it is also important to consider the uncertainties involved in the data collection, the representative characteristics (both spatial and spectral) of the measurement site, and the methods and models involved in deriving the final data product.

The application of the FRM concept to an output requires an understanding of the measurement procedure and the processing required. Importantly, it also requires an understanding of the ingrained assumptions in both. A general-purpose procedure (based on Fahy et al., 2022) to build this understanding can be gained by following these steps:

1. Identify the quantity being retrieved and define it as explicitly as possible.
 - a. Determine if it can be measured in the field (or at all), and if not,
 - b. identify what is being measured in the field and the relationship/model that converts between the measured quantity and the desired quantity.
 - c. For example, do not define “surface reflectance”, use a term which explicitly considers the angular geometries of the incoming and outgoing fluxes. Several of the surface reflectance quantities (e.g. BRDF) cannot be measured explicitly.
2. Identify the input values required to produce the quantity defined in 1.
 - a. Repeat step 1 for these quantities.
3. Identify the instrument/s used to measure the quantities in step 2.
4. Gather the calibration and characterisation information for the instruments identified in step 3. If possible, repeat step 3 for each instrument.
5. Gather the processing instructions to convert the input values identified in step 2 to the quantity defined in step 1.
6. Draw out the processing chain from the perspective of the inputs mapping to the outputs.
7. Go through each link in the processing chain drawn out in step 6 and create a new processing chain where each link is the mathematical function required to convert the

input into that link into the output from that link. This may require replacing 1 link with multiple links.

- a. If the link is a black box (i.e. it is not feasible to get the mathematical formulation for that step), include the link as $output = f(input_1, input_2, \dots, input_n)$.
 - b. For all links, include a “+ 0” term at the end to account for remaining uncharacterised assumptions.
8. Go through each link identified in step 7 and identify if there are any ingrained assumptions at that step. This is known as a measurement model. An example would be where an instrument is assumed to be levelled, introducing a $\cos(x)$ term to account for the orientation of the sensor.
 9. Repeat steps 1-8 for each input until the outermost input is reached.

From this, an uncertainty tree diagram (or flow diagram) can be produced, which then creates a clear series of connections between measurement inputs/outputs aids in developing the uncertainty propagation equation. This type of assessment can facilitate understanding of correlation structures due to linkages in the diagram. An example of an uncertainty propagation tree can be found in Figure 17, with each input area given a different colour and the relationship between various inputs given some clarity. It should be noted that this diagram defines the overall relationship, the exact parameters will be dynamic, probably changing with every new measurement made.

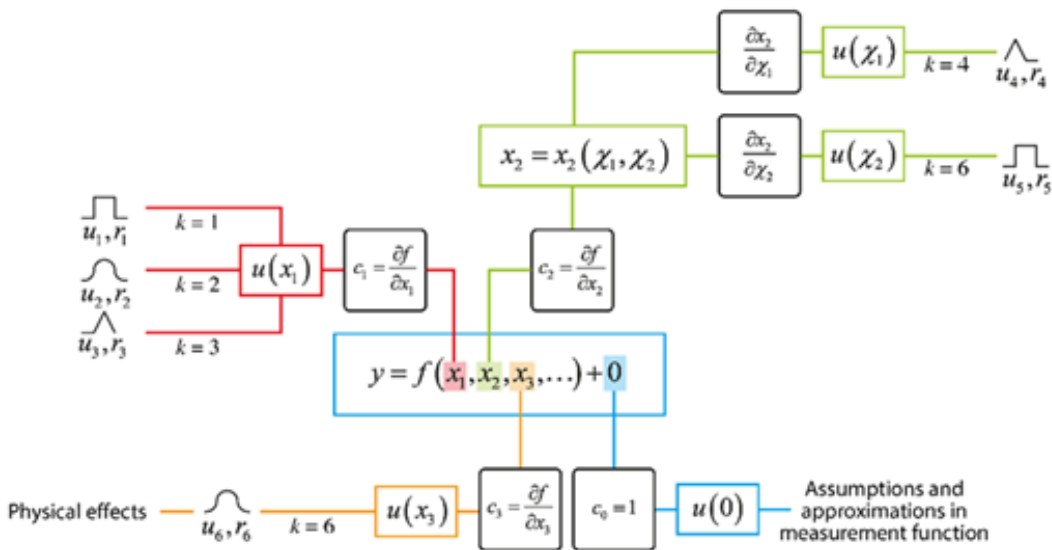


Figure 17. Example uncertainty propagation tree (FIDUCEO, “Example Uncertainty Analysis Tree.” <https://research.reading.ac.uk/fiduceo/fcdrs/theoretical-basis-2/2-defining-uncertainty-effects/> (accessed Nov. 14, 2024)).

The next step in the process involves compiling an effects table for each of the inputs identified in the propagation tree diagram. The aim of the effects table is to identify the contribution of the

uncertainty (the magnitude), to define the probability distribution of the uncertainty and to determine if there is any correlation or structure to consider. Typically, the effects table will have the following headers (Fahy et al., 2022):

- Unique identifier of the effect (name/code).
- Affected parameters in the measurement model.
- Magnitude of the uncertainty.
- Sensitivity coefficients w.r.t. each parameter.
- Probability distribution.
- Correlation and structure.
- An indicator of the maturity of the uncertainty evaluation (e.g. using judgement, calculated through experimentation, calibration certificate, etc.).

An example of an effects table can be found in Figure 18.

Variable	Description	Partial Derivative	Effects
C_E	Earth Scene Counts	$\frac{\partial L_E}{\partial C_E} = a_1$	Noise, Non-linearity
$\overline{C_{BB1}}$	BB1 counts	$\frac{\partial L_E}{\partial \overline{C_{BB1}}} = a_1 \frac{(\overline{C_{BB1}} - C_E)}{(\overline{C_{BB1}} - \overline{C_{BB2}})}$	Noise, Non-linearity
$\overline{C_{BB2}}$	BB2 counts	$\frac{\partial L_E}{\partial \overline{C_{BB2}}} = a_1 \frac{(C_E - \overline{C_{BB1}})}{(\overline{C_{BB1}} - \overline{C_{BB2}})}$	Noise, Non-linearity
L_{BB1}	BB1 radiance	$\frac{\partial L_E}{\partial L_{BB1}} = \frac{(C_E - \overline{C_{BB1}})}{(\overline{C_{BB1}} - \overline{C_{BB2}})}$	BB Temperature Cavity Emissivity Spectral Response
L_{BB2}	BB2 radiance	$\frac{\partial L_E}{\partial L_{BB2}} = \frac{(\overline{C_{BB1}} - C_E)}{(\overline{C_{BB1}} - \overline{C_{BB2}})}$	BB Temperature Cavity Emissivity Spectral Response

Figure 18. Example of an effects table from Smith et al., (2021) giving the table of primary contributions to uncertainties in the Sea and Land Surface Temperature Radiometer (SLSTR) calibration model.

With this information gathered, it is possible to begin propagating the uncertainty via one of two main methods:

- Recreating the processing chain based on the measurement model and running a Monte Carlo (MC) analysis using the information contained in the effects tables to characterise the input parameter distributions,
- Using the law of propagation of uncertainties. This involves calculating the partial differentials, at each step, of the inputs in relation to the output and combining these with the uncertainties given in the effects tables. Equation 8 describes the relationship.

$$u^2(\mathbf{y}) = \sum_{i=1}^N \left(\frac{\partial f}{\partial x_i} \right)^2 u^2(x_i) + 2 \sum_{i=1}^{N-1} \sum_{j=i+1}^N \frac{\partial f}{\partial x_i} \frac{\partial f}{\partial x_j} u(x_i, x_j) \quad 8$$

Where $u(y)$ gives the uncertainty in the output parameter, $\frac{\partial f}{\partial x_i}$ give the sensitivity coefficients, and $u(x_i, x_j)$ give the covariance between input parameters. Where more complicated correlation structures exist, other methods such as the CURUC approach given in Merchant et al., (2019) can be used instead.

There are positives and negatives to both methods. For example, MC analysis is simple to implement, considers the full correlation structures present, and handles non-linear processes. However, it often requires many samples and therefore a significant amount of computational power. For time-consuming processes, this is either not feasible or significant computation facilities are required to complete it. The analytical method (law of propagation of uncertainties) offers a way of computing the uncertainty in a single equation, but it does require calculation of the derivatives, quantitative understanding of the correlation values, and additional terms to describe non-linear functions.

The selection of either method will depend on the complexity of the measurement model and correlation structures, and the time required to complete a single iteration of the processing.

With respect to in-field measurements, the categories of uncertainty sources are listed in the subsections below so that users can consider assessing these in their uncertainty budgets.

6.8.1 Radiometry

A non-exhaustive list of potential uncertainties in radiometric measurements is given in Table 1.

Table 1. Potential instrument effects and reported uncertainties associated with these effects. Where effects are too dependent on multiple factors or instruments, or not reported, the values are omitted. Note that not all values are derived from UAV-based imagers and there are limited published sources. References to the National Environmental Research Council (NERC) Field Spectroscopy Facility (FSF) reflect an unpublished document titled “Calibration and Characterisation of Hyperspectral Imagers” produced by Robbie Ramsey (NERC FSF) w.r.t. to characterisation of a Headwall Co-Aligned imager. NPL characterisation refers to unpublished data collected at the National Physical Laboratory with a Headwall Co-Aligned imager. The instrument used in Baumgartner et al., (2012), Lenhard et al., (2015), and Baumgartner (2021) is the NEO HySpex VNIR-1600 and SWIR-320m-e which is an airborne imaging spectrometer. The instrument used in Barreto et al., (2019) is the Headwall Nano imager. The values given aim to capture the range of expected values for users; the uncertainty in any scenario may be different and users are encouraged to do the full suite of instrument characterisation.

Source	Typical uncertainty (2σ)	Notes
Radiometric calibration	2% – 5% (400 nm – 1000 nm)	NERC FSF; Lenhard et al., (2015); Baumgartner (2021)
Radiometric resolution	0.5 DN	Bit dependent
Detector linearity	3% – 15% (if uncorrected)	Baumgartner (2021)
Dark current	0.1 DN – 0.6 DN (400 nm – 1000 nm).	Lenhard et al., (2015); Barreto et al., (2019). Values from various

		sources but can be easily determined from data.
Spectral calibration	< 1.5 nm (400 nm – 1000 nm)	Baumgartner et al., (2012) Lenhard et al., (2015)
Spectral resolution	-	-
Spectral response	-	-
Signal-to-noise / noise-equivalent spectral radiance	-	-
Temperature dependence	-	-
Stray light	-	Instrument dependent
Aging	-	Usage dependent
Repeatability	< 2%	NPL
Dead pixels	-	-
Bright pixels	-	-
Saturation	-	-
Uniformity (imaging instruments)	Up to 20% (uncorrected); 2% corrected	NPL; Lenhard et al., (2015)
Spectral smile (imaging instruments)	< 1.5 nm	Baumgartner et al., (2012); Baumgartner (2021); Lenhard et al., (2015). Likely uncertainty if uncorrected.
Keystone (imaging instruments)	33% – 56% (of a pixel)	Baumgartner et al., (2012); Baumgartner (2021) (impact on reflectance depends on scene).

- Potential in field effects:
 - Clear-sky models
 - Bidirectional Reflectance Distribution Function models
 - Sensor optics dirt/debris accumulation
 - Instrument levelling (permanent)
 - Temporal stability of illumination conditions (i.e. within scan sequence)
 - Atmospheric effects
- Potential site characterisation effects:
 - Spatial representativeness

- Spatial homogeneity
- Adjacency
- Temporal consistency (in above factors)
- Potential platform effects:
 - Mounting platform (present in field of view, shadowing, etc.)
- Potential campaign-only effects:
 - User competency
 - Instrument levelling (repeated)
 - User within field of view / shadowing / adjacency
 - Levelling of calibration target (e.g. for reflectance)
 - Sampling representativeness
 - Sampling number
 - Instrument stability (e.g. warm-up)
 - Instrument gain setting
- Potential processing effects:
 - Spectral resampling
 - Spectral convolution
 - Spatial resampling
 - Orthorectification

6.8.2 Geometry

Potential geometric uncertainty sources include:

- Geolocation
- Platform motion
- Inertial measurement unit uncertainties (e.g. for UAVs)
- Lens distortions (radial, barrel, pincushion, etc.)
- Field of view characterisation
- Instrument orientation
- Band co-registration

6.8.3 Image quality

Potential image quality uncertainty sources include:

- Radiometric striping
- Sharpness / blur
- Orthorectification
- Image gaps (e.g. UAV images)
- Terrain distortions
- Image stitching (e.g. UAV images)

6.8.4 Simulations

Potential simulation uncertainty sources include:

- Number of simulation iterations (for Monte Carlo)
- Characterisation of simulation parameters
- Ingrained assumptions
- Calibration parameters (for empirical / semi-empirical models)

7 STRATEGY FOR SR VALIDATION BASED ON UAV SR MEASUREMENTS

The validation of remote sensing products can be performed quantitatively by comparing against independent reference datasets. The reference dataset is required to have an uncertainty lower than the product being validated. In the case of the MODIS Level-2 (L2) atmospheric correction algorithm, the uncertainties of BOA surface reflectance (MOD09) retrievals are typically $0.05\rho_{ref} + 0.005$ (Vermote and Kotchenova 2008) (at 68% coverage factor with ρ_{ref} being the measured (reference) surface reflectance. This is the same for the HCRF product derived from Sentinel-2 (Clerc et al., 2019). These requirements can be utilised when there are uncertainty estimates on the satellite and in situ quantities. In the simplest case where no requirement on the uncertainty exists, conformity or agreement between the two measurements can be calculated with (following Fahy et al., 2022):

$$E_N = \frac{|R_s - R_r|}{k\sqrt{u^2(R_s) + u^2(R_r) + u^2(R_c)}} \quad 9$$

And occurs when $E_N < 1$, where R_s is the reflectance factor of the satellite product, R_r is the reflectance factor of the reference (UAS), $u(x)$ gives the uncertainty of each of the effects, with $u(R_c)$ being the uncertainty associated with the comparison conditions, and k gives the coverage interval. However, when there are requirements placed on the uncertainty level, conformity exists when:

$$|R_s - R_r| + k\sqrt{u^2(R_s) + u^2(R_r) + u^2(R_c)} < k\gamma \quad 10$$

And:

$$|R_s - R_r| - k\sqrt{u^2(R_s) + u^2(R_r) + u^2(R_c)} > -k\gamma \quad 11$$

Where γ is the maximum uncertainty level. This approach requires that the uncertainties on the satellite and validation data have been calculated correctly, accounting for all necessary inputs. If $E_N > 1$ this suggests that the uncertainties have not included all the necessary components, the magnitude of those uncertainties has not been assessed correctly, or correlation effects have not been accounted for correctly. As such, it offers a diagnostic tool with which to revisit the measurement and processing procedures of the satellite product (and in situ). This method of validation is *highly recommended, and required*, where conformity to the FRM guidelines is needed according to Goryl et al., (2023) [R23].

A common issue is that satellite surface reflectance products do not always have uncertainty estimates. Where this is the case, it is imperative that in situ estimates have a fully characterised uncertainty. A further recommendation (aimed at satellite surface reflectance product producers) is that uncertainties should, by default, be provided on a per-pixel basis on all new products. For pre-existing products, effort should be expended to provide this information based on vicarious calibration activities over representative sites or, when future missions such as the

Traceable Radiometry Underpinning Terrestrial- and Helio- Studies (TRUTHS) are launched, by cross-calibration. Propagating the uncertainty through to the surface reflectance product would then be required. In the absence of satellite surface reflectance product uncertainty, this uncertainty should not be treated as if it is perfectly known (i.e. $u(R_s) = 0$), instead a conservative best estimate of the per-pixel uncertainty should be made and stated clearly in the analysis, along with justification for that value.

8 CONCLUSIONS

This document provides recommendations for good practices in the validation of satellite surface reflectance products using UAVs. The recommendations described are not intended to encompass all applications of UAV-based hyperspectral measurements, however, they may prove useful in many cases. The creation of the document is based on two intercomparison exercises in Barrax, Spain in 2022 and Calperum, Australia in 2024, user workshops following these campaigns, and input from the international scientific community (listed as co-authors). This document is intended to address the implementation of good practice protocols in this area for the purposes of satellite validation and simultaneously advance the validation maturity of surface reflectance data products.

UAV-based sensing, in the context of satellite validation and with respect to field spectrometers, offers increased sampling of the surface, shorter time durations for the measurements, eliminates trampling and other interference at the surface, provides a viewing geometry closer to the satellite pixel, and provides access to harder to measure surfaces (e.g. tall trees, water, etc.). These reasons will be driving the increased uptake of UAV-based hyperspectral remote sensing.

This document draws upon CEOS-FRM (Goryl et al., 2023) which provide requirements on validation data for satellite products. Based on this, there is a distinction between a requirement and a recommendation, where the former is stricter. For example, to reach CEOS-FRM status, the calculation of the uncertainties associated with the validation data is a requirement, as is up-to-date calibration information (for reduced uncertainty, traceability, etc.), since these are core tenets. However, recommendations are different in that they represent the community agreed good practice but don't directly deal with components of CEOS-FRM.

The pace of improvement in UAV and payload (hyperspectral/LiDAR/GNSS/IMU) technology, and the rate of change in regulation, means that the good practices and recommendations contained in this document will need periodic review to ensure that it reflects these future advances in technology, processing and comparison methods.

9 References

- Aasen, H., Honkevaara, E., Lucieer, A., & Zarco-Tejada, P.J. (2018) Quantitative Remote Sensing at Ultra-High Resolution with UAV Spectroscopy: A Review of Sensor Technology, Measurement Procedures, and Data Correction Workflows. *Remote Sensing*, **10**(6): 1091.
- Anderson, J., Dugan, K., & MacArthur, A. (2011) On the Reproducibility of Field-Measured Reflectance Factors in the Context of Vegetation Studies. *Remote Sensing of Environment*, **115**: 1893–1905.
- Arroyo-Mora, J.P., Kalacska, M., Inamdar, D., Soffer, R., Lucanus, O., Gorman, J., Naprstek, T., Schaaf, E.S., Ifimov, G., Elmer, K., & Leblanc, G. (2019) Implementation of a UAV–Hyperspectral Pushbroom Imager for Ecological Monitoring. *Drones*, **3**(1): 12.
- Arroyo-Mora, J.P., Kalacska, M., Lucanus, O., Laliberté, R., Chen, Y., Gorman, J., Marion, A., Coulas, L., Barber, H., Borshchova, I., Soffer, R., Leblanc, G., Lavigne, D., Girard, L., & Bérubé, M. (2023) Development of a Novel Implementation of a Remotely Piloted Aircraft System over 25 kg for Hyperspectral Payloads. *Drones*, **7**: 652.
- Barreto, M. A. P., Johansen, K., Angel, Y., & McCabe, M. F. (2019) Radiometric assessment of a UAV-based push-broom hyperspectral camera. *Sensors*, **19**(21): 4699.
- Baumgartner, A., Gege, P., Köhler, C., Lenhard, K., & Schwarzmaier, T. (2012) Characterisation methods for the hyperspectral sensor HySpex at DLR's calibration home base. *Proceedings of SPIE 8533, Sensors, Systems, and Next-Generation Satellites XVI*, 85331H, (19th November 2012).
- Baumgartner, A. (2022) *Traceable Imaging Spectrometer Calibration and Transformation of Geometric and Spectral Pixel Properties*. PhD Thesis, Universität Osnabrück. Available at: <https://osnadocs.uni-osnabrueck.de/handle/ds-202202076056> (Accessed: 06/12/2024).
- Behnert, I., Deadman, A., Fox, N., Harris, P., Gürol, S., Özen, H., Bachmann, M., Boucher, Y., and Lachéradé, S. (2011) Measurement report CEOS WGCV pilot comparison of techniques and instruments used for the vicarious calibration of land surface imaging through a ground reference standard test site 2009. NPL Report OP5, available from: <https://eprintspublications.npl.co.uk/4933/1/OP5.pdf> [Accessed: 11/04/2024].
- Berry, J. K., Heimes, F. J., & Smith, J. A. (1978) A portable instrument for simultaneous recording of scene composition and spectral reflectance. *Optical Engineering*, **17**: 143–146.
- Bialek, A., Greenwell, C., Lamare, M., Marcq, S., Lacherade, S., and Meygret, A. (2016) Namibia Field Campaign Technical Note. NPL Report.
- Bouvet, M., Thome, K., Berthelot, B., Bialek, A., Czaplá-Myers, J., Fox, N. P., Goryl, P., Henry, P., Ma, L., Marcq, S., Meygret, A., Wenny, B. N. and Woolliams, E. R. (2019) RadCalNet: A Radiometric

Calibration Network for Earth Observing Imagers Operating in the Visible to Shortwave Infrared Spectral Range. *Remote Sensing*, **11**(20): 2401.

Brach, E. J., Poirier, P., Desjardins, R. L., & Lord, D. (1983) Multispectral radiometer to measure crop canopy characteristics. *Review of Scientific Instruments*, **54**: 493–500.

Burggraaff, O. (2020) Biases from incorrect reflectance convolution. *Optics Express*, **28**(9): 13801 – 13816.

Byrne, G., Walsh, A., Thankappan, M., Broomhall, M., & Hay, E. 2021. DEA Analysis Ready Data Phase 1 Validation Project: Data Summary. Geoscience Australia, Canberra. <https://doi.org/10.26186/145101>.

Campbell, S., Lovell, J., Jupp, D.L.B., Graetz, R.D., & Byrne, G. (2001). The Lake Frome field campaign in support of Hyperion instrument calibration and validation. In IGARSS 2001 Scanning the Present and Resolving the Future, IEEE 2001 International Geoscience and Remote Sensing Symposium Proceedings 6. 2593–2595 vol.6. 10.1109/IGARSS.2001.978099.

Chen, X., & Vierling, L. A. (2006) Spectral mixture analyses of hyperspectral data acquired using a tethered balloon. *Remote Sensing of Environment*, **103**: 338–350.

Clerc, S., Devignot, O., Pessiot, L., Team, M.P.C., 2019. S2 MPC. Level 2A Data Quality Report. In: Technical Report S2-PDGS-MPC-L2ADQR. European Space Agency. <https://sentinel.esa.int/documents/247904/685211/Sentinel-2-L2A-Data-Quality-Report> [Accessed: 06/03/2024].

Cooksey, C.C., Allen, D.W., Tsai, B.K. & Yoon, H.W. (2015). Establishment and application of the 0/45 reflectance factor scale over the shortwave infrared. *Applied Optics*, **54**(10): 3064–3071.

Doelling, D., Helder, D., Schott, J., Stone, T., & Pinto, C. (2017) Vicarious Calibration and Validation. In Reference Module in Earth Systems and Environmental Sciences; Elsevier: Amsterdam, The Netherlands; pp. 475–518.

Doxani, G., Vermote, E., Roger, J.-C., Gascon, F., Adriaensen, S., Frantz, D., Hagolle, O., Hollstein, A., Kirches, G., Li, F., Louis, J., Mangin, A., Pahlevan, N., Pflug, B. & Vanhellefont, Q. (2018) Atmospheric correction inter-comparison exercise. *Remote Sensing*, **10**(2): 352.

Doxaran, D., ElKilani, B., Corizzi, A., Goyens, C. (2023) Validation of satellite-derived water-leaving reflectance in contrasted French coastal waters based on HYPERNETS field measurements. *Frontiers in Remote Sensing*, **4**: 1290110.

Dreier, A., Janßen, J., Kuhlmann, H., Klingbeil, L. (2021) Quality Analysis of Direct Georeferencing in Aspects of Absolute Accuracy and Precision for a UAV-Based Laser Scanning System. *Remote Sensing*, **13**(18): 3564.

Fahy, J., Fox, N., Gardiner, T., Green, P., Hunt, S., Mittaz, J., Mota, B., De Vis, P. & Woolliams, E. (2022) General Guidance on a Metrological Approach to Fundamental Data Records (FDR), Thematic Data Products (TDP) and Fiducial Reference Measurements (FRM) – Uncertainty Analysis Process. London: NPL Management Ltd. Available from: http://qa4eo.org/docs/3_Process_Document.pdf [Accessed: 06/03/2024].

Gamon, J.A., Rahman, A.F., Dungan, J.L., Schildhauer, M. & Huemmrich, K.F. (2006a) Spectral Network (SpecNet) – What is it and why do we need it? *Remote Sensing of Environment*, **103**: 227–235.

Gamon, J. A., Cheng, Y., Claudio, H., MacKinney, L., & Sims, D. A. (2006b) A mobile tram system for systematic sampling of ecosystem optical properties. *Remote Sensing of Environment*, **103**: 246–254.

GCOS-244/272. “GCOS 2022 Implementation Plan”. *Global Climate Observing System*. (2022)

GCOS-245. “The 2022 GCOS ECVs Requirements”. *Global Climate Observing System*. (2022)

Goetz, A. F. H. (1975) “Appendix E: Portable field reflectance spectrometer” in “Application of ERTS images and image processing to regional geologic problems and geologic mapping in Northern Arizona”. JPL Technical Report 32–1597 (pp. 183–188).

Goetz, A. F. H. (1987). The portable instant display and analysis spectrometer (PIDAS). Proceedings of the Third Airborne Imaging Spectrometer Data Analysis Workshop, Vol. 87–30. (pp. 8–17): JPL Publication.

Goetz, A.F.H. (2009) Three decades of hyperspectral remote sensing of the Earth: A personal view, *Remote Sensing of Environment*, **113**: S5–S16

Goetz, A.F.H. (2012) Making Accurate Field Spectral Reflectance Measurements. ASD Inc Technical Note, available from: <https://www.materials-talks.com/wp-content/uploads/2018/01/Making-Accurate-Field-Spectral-Reflectance-Measurements-LR.pdf> [Accessed: 06/03/2024].

González Vilas, L., Brando, V.E., Concha, J.A., Goyens, C., Dogliotti, A.I., Doxaran, D., Dille, A., & Van der Zande, D. (2024) Validation of satellite water products based on Hypernets in situ data using a Match-up Database File (MDB) structure. *Frontiers in Remote Sensing*, **5**: 1330317.

Goryl, P., Fox, N., Donlon, C., & Castracane, P. (2023) Fiducial Reference Measurements (FRMs): What Are They? *Remote Sensing*, **15**(20): 5017.

Goyens, C., De Vis, P., & Hunt, S.E. (2021) Automated Generation of Hyperspectral Fiducial Reference Measurements of Water and Land Surface Reflectance for the Hypernets Networks, *2021 IEEE International Geoscience and Remote Sensing Symposium IGARSS*, 11–16 July 2021, DOI: 10.1109/IGARSS47720.2021.9553738.

Gras, J.-P., Barthès, B.G., Mahaut, B., & Trupin, S. (2014) Best practices for obtaining and processing field visible and near infrared (VNIR) spectra of topsoils. *Geoderma*, **214–215**: 126–134.

Held, A., Phinn, S., Soto-Berelov, M. & Jones, S. (Eds) (2015) "AusCover good practice guidelines: A technical handbook supporting calibration and validation activities of remotely sensed data products." *TERN AusCover* 352.

Honkavaara, E. and Khoramshahi, E. (2018) Radiometric Correction of Close-Range Spectral Image Blocks Captured Using an Unmanned Aerial Vehicle with a Radiometric Block Adjustment. *Remote Sensing*, **10**(2): 256.

Hueni, A. Damm, A. Kneubuler, M. Schlapfer, D. & Schaepman, M. (2017) Field Airborne Spectroscopy Cross Validation: Some Considerations. *IEEE Journal of Selected topics in Applied Earth Observation and Remote Sensing*, **10**(3): 1117–1135.

Hueni, A. & Bialek, A. (2017) Cause, effect, and correction of field spectroradiometer interchannel radiometric steps. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, **10**(4): 1542–1551.

Inamdar, D., Kalacska, M., Leblanc, G. and Arroyo-Mora, J. P. (2020) Characterising and Mitigating Sensor Generated Spatial Correlations in Airborne Hyperspectral Imaging Data. *Remote Sensing*, **12**(4): 641.

Inamdar, D., Kalacska, M., Arroyo-Mora, J.P., & Leblanc, G. (2021) The directly-georeferenced hyperspectral point cloud: preserving the integrity of hyperspectral imaging data. *Frontiers in Remote Sensing*, **2**: 675323.

Inamdar, D., Soffer, R., Kalacska, M., & Naprstek, T. (2023) Spatial Coverage Map and Resampling Error Assessment for Hyperspectral Imaging Data, doi: 10.5683/SP3/EO8LM8, Borealis, V1.

Joint Committee for Guides and Measures (JCGM) (2008) Evaluation of measurement data - guide to the expression of uncertainty in measurement. Technical report: 100. Available from: <https://www.bipm.org/en/committees/jc/jcgm/publications> [Accessed: 06/03/2024].

Joint Committee for Guides in Metrology (JCGM) (2021) International Vocabulary of Metrology, 4th Edition. URL: https://www.bipm.org/documents/20126/54295284/VIM4_CD_210111c.pdf/a57419b7-790f-2cca-f7c9-25d54d049bf6 [accessed 24/02/2024].

Kalacska, M., Arroyo-Mora, J.P., Soffer, R., & Elmer, K. (2019) ASD FieldSpec3 field measurement protocols, *Protocols.io*, doi: 10.17504/protocols.io.qu7dwzn.

Kalacska, M., Arroyo-Mora, J.P., Lucanus O., Sousa, L., Pereira, T. & Vieira, T (2020) Deciphering the many maps of the Xingu River Basin – an assessment of land cover classifications at multiple scales, *Proceedings of the Academy of Natural Sciences of Philadelphia*, **166**(1): 1–55.

- Kruse, F.A., Lefkoff, A.B., Boardman, J.W., Heidebrecht, K.B., Shapiro, A.T., Barloon, P.J. & Goetz, A.F.H. (1993) The spectral image processing system (SIPS) - interactive visualisation and analysis of imaging spectrometer data. *Remote Sensing of Environment*, **44**(2): 145–163.
- Lenhard, K., Baumgartner, A., & Schwarzmaier, T. (2015) Independent laboratory characterisation of NEO HySpex imaging spectrometers VNIR-1600 and SWIR-320m-e. *IEEE Transactions on Geoscience and Remote Sensing*, **53**(4): 1828–1841.
- Leuning, R., Hughes, D., Daniel, P., Coops, N. C., & Newnham, G. (2006) A multi-angle spectrometer for automatic measurement of plant canopy reflectance spectra. *Remote Sensing of Environment*, **103**: 236–245.
- Lewis, P., & Barnsley, M. J. (1994). Influence of the sky radiance distribution on various formulations of the earth surface albedo. 6th International Symposium on Physical Measurements and Signatures in Remote Sensing, ISPRS (pp. 707–715). Val d'Isere, France: CNES.
- Lewis, A., Oliver, S., Lymburner, L., Evans, B., Wyborn, L., Mueller, N., Raevski, G., Hooke, J., Woodcock, R., Sixsmith, J., Wu, W., Tan, P., Li, F., Killough, B., Minchin, S., Roberts, D., Ayers, D., Bala, B., Dwyer, J., Dekker, A., Dhua, T., Hicks, A., Ip, A., Purss, M., Richards, C., Sagar, S., Trenhamb, C., Wang, P. & Wang, L.-W. (2017) The Australian Geoscience Data Cube - Foundations and lessons learned. *Remote Sensing of Environment*, **202**: 276–292.
- Lucht, W., Schaaf, C. B., & Strahler, A. H. (2000). An algorithm for the retrieval of albedo from space using semiempirical BRDF models. *IEEE Transactions on Geoscience and Remote Sensing*, **38**: 977–998.
- Lucieer, A., Malenovský, Z., Veness, T., & Wallace, L. (2014) HyperUAS – Imaging spectroscopy from a multi-rotor unmanned aircraft system. *Journal of Field Robotics*, **31**(4): 571–590.
- Malthus, T.J., Karpouzli, E., Thankappan, M., Dekker, A., Smith, C. (2014) An audit of satellite calibration and validation facilities and activities in Australia. Report prepared for the Department of Industry Space Coordination Office. CSIRO Future Science Platform in Earth Observation Informatics, Canberra Australia. ISBN 978-1-4863-0490-5.
- Malthus, T., Ong, C., Lau, I., Fearn, P., Byrne, G., Thankappan, M., Chisholm, L., Suarez Barranco, L., Clarke, K., Scarth, P., & Phinn, S. (2019a) A community approach to standardised validation of surface reflectance data – A technical handbook to support the collection of field reflectance data. CSIRO Centre for Earth Observation, Brisbane, Australia; ISBN 978-1-4863-0991-7, doi: 10.25919/5c9d0ba9e9c12.
- Malthus, T., Ong, C., Lau, I., & Fearn, P. (2019b). A community approach to the standardised validation of surface reflectance data – Final Report. CSIRO Centre for Earth Observation, Brisbane, Australia.
- Merchant, C.J., Holl, G., Mittaz, J.P.D., & Woolliams, E.R. (2019) Radiance uncertainty characterisation to facilitate climate data record creation. *Remote Sensing*, **11**(474): 17.

- Milton, E. J., Blackburn, G. A., Rollin, E. M., & Danson, F. M. (1994) Measurement of the spectral directional reflectance of forest canopies: A review of methods and a practical application. *Remote Sensing Reviews*, **10**: 285–308.
- Milton, E.J., Schaepman, M.E., Anderson, K., Kneubühler, M., & Fox, N. (2009) Progress in field spectroscopy, *Remote Sensing of Environment*, **113**: S92–S109.
- Morris, H., Sinclair, M., De Vis, P. and Bialek, A. (2024) Utilising LANDHYPERNETS data products over a deciduous broadleaf forest to validate Sentinel-2 and Landsat surface reflectance products. *Frontiers in Remote Sensing*, **5**: 1322760.
- Naprstek, T., & Inamdar, D. (2022) HyPlanT - a standalone flight planning application for pushbroom hyperspectral imagers, doi: 10.5683/SP3/3QKALR, Borealis, V1.
- Nevalainen, O., Honkavaara, E., Tuominen, S., Viljanen, N., Hakala, T., Yu, X., Hyyppä, J., Saari, H., Pölonen, I., Imai, N., et al., (2017) Individual Tree Detection and Classification with UAV-Based Photogrammetric Point Clouds and Hyperspectral Imaging. *Remote Sensing*, **9**(3): 185.
- Nicodemus, F.E., Richmond, J.C., Hsia, J.J., Ginsberg, I.W. & Limperis, T. (1977). Geometrical considerations and nomenclature for reflectance. Washington, DC: National Bureau of Standards, US Department of Commerce.
- Niro, F., Goryl, P., Dransfeld, S., Boccia, V., Gascon, F., Adams, J., Themann, B., Scifoni, S., & Doxani, G. (2021) European Space Agency (ESA) calibration/validation strategy for optical land-imaging satellites and pathway towards interoperability. *Remote Sensing*, **13**(15): 3003.
- Ong, C., Malthus, T., Lau, I.C., Thankappan, M. & Byrne, G. (2018) The Development of a Standardised Validation Approach for Surface Reflectance Data, IGARSS 2018 - 2018 *IEEE International Geoscience and Remote Sensing Symposium*, Valencia, Spain, pp. 6456–6459.
- Origo, N., Gorroño, J., Ryder, J., Nightingale, J. & Bialek, A. (2020) Fiducial Reference Measurements for validation of Sentinel-2 and Proba-V surface reflectance products. *Remote Sensing of Environment*, **241**: 111690.
- Origo, N., Morris, H., Randall, C., Sinclair, M., Morrone, R., MacClellan, C., Dash, J., Brown, L., Camacho, F., Martínez, E., Sánchez-Zapero, J. & Boccia, V. (in preparation) Validation of Sentinel-2 Surface Reflectance using a UAV-mounted Hyperspectral Imager. *Remote Sensing*.
- Penndorf, R. (1956) Luminous and spectral reflectance as well as colours of natural objects. U.S. Air Force Cambridge Research Center, Bedford, Massachusetts.
- Pfeifer, N., Glira, P., & Briese, C. (2012) Direct georeferencing with on board navigation components of light weight UAV platforms. *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*: **39-B7**, XXII ISPRS Congress, Melbourne, Australia.

Pflug, B, Bachmann, M. De los Reyes, R., Marshall, D., Holzwarth, S., Gorrone, J. and Meygret, A., 2023, Ground based validation over land of the passive optical remote sensing data, LPVE 2023, ESA-ESRIN, Frascati, Italy.

RadCalNet Technical Working Group (2018) RadCalNet Guidance: Site Characterisation. CEOS Reference: QA4EO-WGCV-RadCalNet-G2-v1, Version 1.

Sánchez-Zapero, J., Martínez-Sánchez, E., Camacho, F., Wang, Z., Carrer, D., Schaaf, C., García-Haro, F.J., Nickeson, J. & Cosh, M. (2023) Surface ALbedo VALidation (SALVAL) Platform: Towards CEOS LPV Validation Stage 4—Application to Three Global Albedo Climate Data Records." *Remote Sensing*, **15**(4): 1081.

Schaepman-Strub, G. Schaepman, M Painter, T. Dangel, S., & Martonchik, J. (2006) Reflectance quantities in optical remote sensing - definitions and case studies. *Remote Sensing of Environment*, **103**: 27–42.

Schiller, S. & Silny, J. (2010). The Specular Array Radiometric Calibration (SPARC) method: a new approach for absolute vicarious calibration in the solar reflective spectrum. *SPIE Optical Engineering + Applications*, **7813**.

Schunke, S., Leroy, V. & Govaerts, Y. (2023) Retrieving BRDFs from UAV-based radiometers for fiducial reference measurements: caveats and recommendations. *Frontiers in Remote Sensing*, **4**: 1285800.

Schweiger, A.K. (2020) Spectral Field Campaigns: Planning and Data Collection. In Cavender-Bares, J., Gamon, J.A., & Townsend, P.A. (eds.) *Remote Sensing of Plant Biodiversity*. Springer, Cham. DOI: 10.1007/978-3-030-33157-3_15.

Smith, D., Hunt, S. E., Etxaluse, M., Peters, D., Nightingale, T., Mittaz, J., Woolliams, E. and Polehampton, E. (2021) Traceability of the Sentinel-3 SLSTR Level-1 Infrared Radiometric Processing. *Remote Sensing*, **13**(3): 374.

Soto-Berelev, M., Jones, S., Farmer, E., & Woodgate, E. (2015) Review of validation standards of biophysical Earth Observation products. In Held, A., Phinn, S., Soto-Berelev, M., & Jones, S. (Eds.), *AusCover Good Practice Guidelines: A Technical Handbook Supporting Calibration and Validation Activities of Remotely Sensed Data Product* (pp. 8-30). Version 1.1. TERN AusCover, ISBN: 978-0-646-94137-0.

Steven, M. D. (2004) Correcting the effects of field of view and varying illumination in spectral measurements of crops. *Precision Agriculture*, **5**: 55–72.

Stöcker, C., Bennett, R., Nex, F., Gerke, M. & Zevenbergen, J. (2017) Review of the Current State of UAV Regulations. *Remote Sensing*, **9**(5): 459.

- Suomalainen, J, Hakala, T., Oliveira, R.A., Markelin, L., Viljanen, N., Näsi, R., & Honkevaara, E. (2018) A Novel Tilt Correction Technique for Irradiance Sensors and Spectrometers On-Board Unmanned Aerial Vehicles. *Remote Sensing*, **10**(12): 2068.
- Suomalainen, J., Oliveira, R.A., Hakala, T., Koivumäki, N., Markelin, L., Näsi, R., & Honkevaara, E. (2021) Direct reflectance transformation methodology for drone-based hyperspectral imaging. *Remote Sensing of Environment*, **266**: 112691.
- Teillet, P. M. Barker, J. L. Markham, B. L., Irish, R. R. Fedosejevs, G., & Storey, J. C. (2001) Radiometric cross-calibration of the Landsat-7 ETM+ and Landsat-5 TM sensors based on tandem data sets. *Remote Sensing of Environment*, **78**: 39–54.
- Thome, K. J. (2001) Absolute radiometric calibration of Landsat 7 ETM+ using the reflectance-based method. *Remote Sensing of Environment*, **78**: 27–38.
- Thome, K., Cattrall, C., D'Amico, J., & Geis, J. (2005). Ground-reference calibration results for Landsat-7 ETM+. In: Butler, J., ed., Earth Observing Systems X, 58820B, Proceedings of SPIE, **5882**, SPIE Bellingham WA. (22 August 2005).
- Transport Canada (2021) Remotely Piloted Aircraft Systems Operational Risk Assessment. Document no: AC 903-001, Issue no: 01. Available from: <https://tc.canada.ca/en/aviation/reference-centre/advisory-circulars/advisory-circular-ac-no-903-001> [accessed: 29/02/2024].
- Trevithick, R. (2015) Field data collection and management for Earth Observation image validation. In Held, A., Phinn, S., Soto-Berelov, M. & Jones, S. (Eds.) *Auscover Good Practice Guidelines: A Technical Handbook Supporting Calibration and Validation Activities of Remotely Sensed Data Product* (pp. 31–54). Version 1.1. TERN AusCover, ISBN: 978-0-646-94137-0.
- Vansteenwegen, D., Ruddick, K., Cattrijsse, A., Vanhellemont, Q., & Beck, M. (2019) The Pan-and-Tilt Hyperspectral Radiometer System (PANTHYR) for Autonomous Satellite Validation Measurements—Prototype Design and Testing, *Remote Sensing*, **11**(11): 1360.
- Vermote, E., Tanré, D., Deuzé, J.L., Herman, M., Morcrette, J.J. (1997) Second simulation of the satellite signal in the solar spectrum, 6s: an overview. *IEEE Transactions on Geoscience and Remote Sensing*, **35**(3): 675–686.
- Vermote, E. & Kotchenova, S. (2008) Atmospheric correction for the monitoring of land surfaces. *Journal of Geophysical Research*, **113**: D23S90.
- Yokoya, N., Miyamura, N., and Iwasaki, A. (2010) Preprocessing of hyperspectral imagery with consideration of smile and keystone properties. *Proceedings of SPIE 7857, Multispectral, Hyperspectral, and Ultraspectral Remote Sensing Technology, Techniques, and Applications III*, **78570B**.

Zarco-Tejada, P.J., González-Dugo, V., & Berni, J.A.J. (2012) Fluorescence, temperature and narrow-band indices acquired from a UAV platform for water stress detection using a micro-hyperspectral imager and a thermal camera. *Remote Sensing of Environment*, **117**: 322–337.