A Novel and Open-Source Illumination Correction for Hyperspectral Digital Outcrop Models

Samuel T. Thiele⬤, Sandra Lorenz⬤, Moritz Kirsch⬤, and Richard Gloaguen⬤

Abstract—**The widespread application of drones and associated miniaturization of imaging sensors has led to an explosion of remote sensing applications with very high spatial and spectral resolutions. The 3-D ultrahigh-resolution digital outcrop models created using drones and oblique imagery from ground-based sensors are now commonly used in the academic and industrial sectors, while the generation of spatially accurate models has been greatly facilitated by the development of computer vision tools, such as structure from motion, and the correction of spectral attributes to achieve material reflectance measurements remains challenging. Following the development of a topographical correction toolbox (mephysto), we now propose a series of new tools that can leverage the detailed geometry captured by digital outcrop models to correct for illumination effects caused by oblique viewing angles and the interaction of light with complex 3-D surfaces. This open-source code is integrated into *hylite*, a python toolbox for the full 3-D processing and fusion of digital outcrop models with hyperspectral imaging data. We validate the performance of our novel method using a case study at an open-pit mine in Tharsis, Spain, and demonstrate the importance of accurate illumination corrections for quantitative spectral analyses. Significantly, we show that commonly applied spectral analysis techniques can yield erroneous results for data corrected using current state-of-the-art approaches. Our proposed method ameliorates many of the issues with these established approaches.**

Index Terms—**Digital outcrop models, geology, hyperspectral imaging, illumination correction.**

I. INTRODUCTION

DIGITAL outcrop models are now widely used in the geosciences, as they allow detailed, quantitative, and objective analyses of complex outcrops [1]–[3]. They greatly improve the characterization of outcrop geometry and structure, allowing accurate mapping of, e.g., fractures (e.g., [4], [5]), faults (e.g., [2]), lithology (e.g., [6]), sedimentary architecture, and (for some minerals) mineralogy (e.g., [7], [8]). These datasets, thus, have widespread applications in geohazard management, geotechnical engineering, oil and gas, and the minerals industry [1], [9], [10].

In addition to geometry, digital outcrop models capture color information, from which variations in lithology and composition can be inferred. Recently, this spectrally limited

information is being supplemented with detailed hyperspectral attributes captured using imaging spectrometers, such that individual data points can contain hundreds of spectral bands [7], [11]–[13]. These "hyperclouds" allow the application of advanced remote sensing algorithms to infer outcrop mineralogy and discriminate between lithologies that may appear similar in the visible part of the spectrum [14], [15].

However, spectral artifacts caused by unevenly illuminated topography (e.g., shading and shadowing) need to be corrected before automated mapping approaches can be accurately applied. Hyperspectral cameras need significantly more light than conventional trichromatic (RGB) frame sensors, so they perform best under well-lit (sunny) conditions, especially in longer wavelength regions, such as the shortwave infrared (SWIR). Thus, unlike for, e.g., photogrammetric surveys (for which overcast conditions are ideal), optimal light conditions for hyperspectral data acquisition will generally also cause significant topographic effects.

A variety of methods have been developed to correct these topographic effects in satellite data, but these are challenging to transfer to digital outcrop models due to their high resolution, complex geometry, and often oblique viewing angle [12]. Particular challenges include steep faces (cliffs), curved outcrops (e.g., open-pit mines), and datasets in which indirect illumination from the sky (skylight) is significant (see Fig. 1).

Physics-based approaches developed for satellite or airborne datasets include the topographic correction as part of a larger atmospheric correction model [16]–[19], while simpler techniques (e.g., c-factor and minnaert; cf. [12]) use statistical methods to correct topographic effects independently of any atmospheric influence.

Unlike satellite applications, sensor-to-target distances tend to be small for digital outcrop models, reducing the atmospheric correction to a comparatively simple adjustment for downwelling light composition. This light is normally characterized using one or more calibration panels located in fully illuminated parts of the scene. At-sensor radiance measured for these panels is compared with their known reflectance spectra to derive an empirical line correction (ELC; [20], [21]).

This ELC will be accurate if the illumination is constant throughout a scene and calibration panels are correctly oriented. However, geometric complexities normally cause shading effects due to: 1) variations in the incidence angle of direct sunlight; 2) specular effects (particularly for smooth or polished areas); 3) self-shadowing when a surface is oriented at $>90°$ from the incident light; and 4) projected shadows
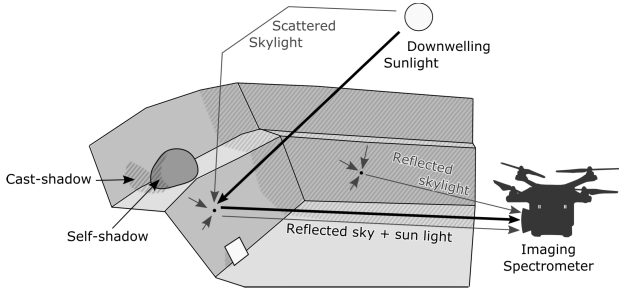
Fig. 1. Conceptual model showing the illumination components in a geometrically complex scene. Shading and shadow effects cause the imaging spectrometer to measure various mixtures of reflected sun and skylight at different locations. These effects are often further complicated by nonnadir viewing directions (as shown).

caused when prominent objects block direct light across the scene.

Shaded areas are rarely completely dark, however, as they are lit by diffuse skylight and reflections from adjacent surfaces. The downwelling light is, thus, a combination of direct illumination from the sun, diffuse skylight, and multiple reflections at any point in a scene. Sunlight is dominant in well-illuminated regions, including calibration panels, while skylight dominates in shaded or shadowed areas.

In the following, we build on the work of [12] to explore challenges associated with topographic correction for digital outcrop models and propose a new joint method that: 1) integrates calibration panel data and information from shadowed areas to separate skylight and sunlight spectra and 2) uses high-resolution geometric data to derive a per-band and per-pixel combined topographic and atmospheric correction. We validate this approach using a digital model of an open-pit mine in the Iberian Pyrite Belt for which hyperspectral data were acquired under both sunny (strong topographic effects) and overcast (minimal topographic effects) conditions. By comparing the two scenes, topographic influence can be independently quantified and used as a baseline to evaluate different correction algorithms.

II. METHODS

State-of-the-art algorithms used for correcting digital outcrop data have been reviewed recently by Jakob *et al.* [12]. These methods all consider only a single light source and, in essence, transform atmospherically corrected per-pixel reflectance $\mathbf{r}_0$ to a topographically corrected value $\mathbf{r}_c$ based on some function of the cosine of the angle between the surface normal vector $\mathbf{n}$ at each pixel and the (downward-pointing) solar illumination vector $\mathbf{s}$.

Formulas for calculating incidence angle from slope and aspect angles can be found in the literature, but, because digital outcrop models store surface normals as an upward-pointing set of cartesian vectors $\mathbf{n}_i$, we find it convenient to work with the cosine of this angle, the dot-product of $\mathbf{s}$ and $\mathbf{n}_i$. This value (α) represents the portion of incident light that would be reflected if we assume the outcrop is a Lambertian reflector illuminated by a single light source

$$\alpha = \max(0, -\mathbf{n_i} \cdot \mathbf{s}). \tag{1}$$

However, in reality, neither of these assumptions holds: the roughness of geological outcrops makes them non-Lambertian reflectors, and downwelling skylight (as well as complex illumination effects caused by light reflected off adjacent objects) means that the measured radiance is a function of multiple light sources, each with distinct irradiance spectra. Research on non-Lambertian reflectors is abundant, allowing us to replace the Lambertian definition of α (1) with, for example, an Oren–Nayar model for diffuse reflection from rough surfaces [22]. This simulates reflection off microfacets oriented with standard deviation σ given the incident light orientation (θ_i, ϕ_i) and viewing angle (θ_r, ϕ_r) to derive an adjusted α value

$$\alpha = \cos \theta_i \cdot (A + B \cdot \max[0, \cos(\phi_i - \phi_r)] \cdot \sin \omega \cdot \tan \beta) \tag{2}$$

where

$$
\begin{aligned}
A &= 1 - 0.5 \frac{\sigma^2}{\sigma^2 + 0.33} \\
B &= 0.45 \frac{\sigma^2}{\sigma^2 + 0.09} \\
\omega &= \max[\theta_i, \theta_r] \\
\beta &= \min[\theta_i, \theta_r].
\end{aligned}
\tag{3}
$$

When roughness (σ) is zero, then this reduces to (1). We use this Oren–Nayar model here as it is a simple approximation for most rock types [23], but a more complex bidirectional reflectance distribution function (BRDF) could easily be substituted if necessary.

The influence of other light sources, especially downwelling skylight, remains a challenge to incorporate into correction algorithms. The skylight received by a point can be modeled using the sky-view factor a, which ranges from 0 to 1 and represents the proportion of skylight relative to a horizontal plane. It can be calculated using a variety of methods [24]–[26], including the portion de ciel visible (PCV) plugin in CloudCompare [27].

Following [28], skylight can be incorporated into a two-light source (sunlight + skylight) model for the relation between material reflectance (**R**) and measured radiance (**r**)

$$r_\lambda = (\alpha \cdot I_\lambda + a \cdot S_\lambda) R_\lambda \tag{4}$$

where I_λ and S_λ are the intensities of incident sunlight and downwelling skylight at wavelength λ, R_λ is the material reflectance at this same wavelength, and r_λ is the reflected radiance. It is straightforward to calculate α and a from the high-resolution geometry captured in a digital outcrop model (cf. Section III-B), allowing (4). to be used as a joint topographic and atmospheric correction if the incident sunlight and diffuse skylight spectra can be estimated or measured.

Assuming that calibration panel orientation is known, either from measurements in the field or based on its projected geometry in the hyperspectral image and known camera orientation, then it is possible to relate these two light sources with measured target radiance by substituting the relevant values into (4)

$$r_\lambda^{\mathrm{panel}} = (\alpha^{\mathrm{panel}} \cdot I_\lambda + a^{\mathrm{panel}} \cdot S_\lambda) R_\lambda^{\mathrm{panel}} \tag{5}$$

and rearranging to derive

$$\alpha^{\text{panel}} \cdot I_\lambda + a^{\text{panel}} \cdot S_\lambda = \frac{r_\lambda^{\text{panel}}}{R_\lambda^{\text{panel}}}. \tag{6}$$

A. Correction With Known Skylight

If the downwelling skylight radiance (S_λ) is known, ideally by including a completely shaded calibration panel or by assuming relatively uniform sky color and sampling sky-pixels visible in the hyperspectral imagery, then the incident sunlight spectra (I_λ) can be directly calculated

$$I_\lambda = \frac{1}{\alpha^{\text{panel}}} \left(\frac{r_\lambda^{\text{panel}}}{R_\lambda^{\text{panel}}} - a^{\text{panel}} \cdot S_\lambda \right) \tag{7}$$

and the joint atmospheric-topographic correction applied by rearranging (4) to calculate per-pixel reflectance

$$R_\lambda = \frac{r_\lambda}{(\alpha \cdot I_\lambda + a \cdot S_\lambda)}. \tag{8}$$

B. Correction With Estimated Skylight

For scenes containing abundant shadows, the skylight spectra can be estimated statistically by assuming that the median reflectance of areas affected by shadows is equal to the overall median reflectance. This assumption is rather limiting but generally holds if shadows are evenly distributed across a scene. Using square brackets to denote the median operator and $\mathbf{R}^{\text{shade}}$ to denote the shaded subset of pixel reflectance ($\mathbf{R}$), we can express this core assumption as

$$[\mathbf{R}_\lambda] \approx \left[\mathbf{R}_\lambda^{\text{shade}} \right]. \tag{9}$$

Shade pixels can be identified using outcrop geometry to: 1) map self-shadowed regions with normal-vectors pointing away from the sun and 2) identify projective shadows using one of the various published geometric or spectral methods (see [29] for a review). By rearranging (4), the general relationship between measured radiance $\mathbf{r}$ and reflectance $\mathbf{R}$ becomes

$$\mathbf{R}_\lambda = \frac{\mathbf{a} \cdot S_\lambda}{\mathbf{r}_\lambda} + \frac{\alpha \cdot I_\lambda}{\mathbf{r}_\lambda} \tag{10}$$

which, for the set of completely shaded pixels where $\alpha = 0$, simplifies to

$$\mathbf{R}_\lambda^{\text{shade}} = \frac{\mathbf{a} \cdot S_\lambda}{\mathbf{r}_\lambda^{\text{shade}}}. \tag{11}$$

Based on the assumption of equal median reflectance (9), we can, thus, derive

$$\left[\frac{\mathbf{a} \cdot S_\lambda}{\mathbf{r}_\lambda^{\text{shade}}} \right] = \left[\frac{\mathbf{a} \cdot S_\lambda}{\mathbf{r}_\lambda} + \frac{\alpha \cdot I_\lambda}{\mathbf{r}_\lambda} \right]. \tag{12}$$

Removing constants from the median operator gives the following relationship between the incident light spectra and diffuse skylight spectra:

$$S_\lambda \left[\frac{\mathbf{a}}{\mathbf{r}_\lambda^{\text{shade}}} \right] = S_\lambda \left[\frac{\mathbf{a}}{\mathbf{r}_\lambda} \right] + I_\lambda \left[\frac{\alpha}{\mathbf{r}_\lambda} \right] \tag{13}$$

$$\therefore I_\lambda = S_\lambda \left(\frac{\left[\frac{\mathbf{a}}{\mathbf{r}_\lambda^{\text{shade}}} \right] - \left[\frac{\mathbf{a}}{\mathbf{r}_\lambda} \right]}{\left[\frac{\alpha}{\mathbf{r}_\lambda} \right]} \right) = S_\lambda \cdot \delta. \tag{14}$$

The term in brackets, which we have replaced with δ in the following equations for clarity, is essentially the normalized reduction in the median radiance of shaded pixels relative to the whole population. It consists entirely of known values, so it can be directly calculated based on the scene geometry and measured hyperspectral radiance data.

Adding additional constraints on the incident light intensity from the calibration panel (7), we can solve for S_λ and I_λ

$$I_\lambda = S_\lambda \cdot \delta = \frac{1}{\alpha^{\text{panel}}} \left(\frac{r_\lambda^{\text{panel}}}{R_\lambda^{\text{panel}}} - a^{\text{panel}} \cdot S_\lambda \right) \tag{15}$$

$$\therefore S_\lambda = \frac{r_\lambda^{\text{panel}}}{R_\lambda^{\text{panel}}} \cdot \frac{1}{a^{\text{panel}} + \delta \cdot \alpha^{\text{panel}}} \tag{16}$$

$$\therefore I_\lambda = \frac{r_\lambda^{\text{panel}}}{R_\lambda^{\text{panel}} \cdot \alpha^{\text{panel}}} \left(1 - \frac{a^{\text{panel}}}{a^{\text{panel}} + \delta \cdot \alpha^{\text{panel}}} \right). \tag{17}$$

These estimated intensities can then be used in (8) to derive corrected reflectance data. Note that zeros (pixels with radiance below detection limit) should be removed from the dataset to avoid infinite values when calculating δ (14).

C. Postprocessing Steps

Regardless of the method used, overcorrections or undercorrections can be caused by geometric errors in the digital outcrop model and misalignments between this and the hyperspectral image. This tends to cause erroneously reflective stripes along with sharp geometric features, such as corners. These areas usually represent a very small fraction of the total dataset, so they can be easily removed by outlier detection methods. The simplest of these is to remove the most/least reflective pixels using a percentile clip: a 1%–99% clip was applied to all the results presented in Section III.
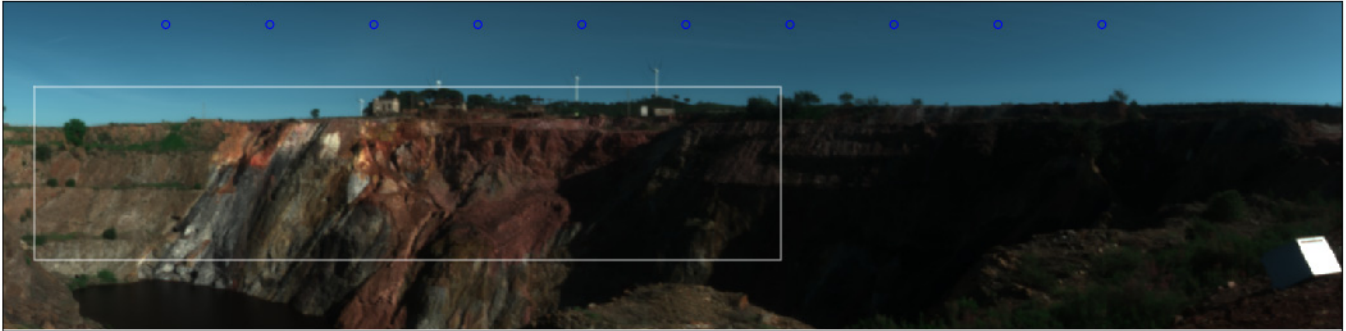
III. VALIDATION STUDY

To validate our novel joint illumination correction method and compare it to the existing state-of-the-art approaches reviewed by Jakob *et al.* [12], we present results from a case study using data from Sierra Bullones (SB), an open-pit mine in the Tharsis mining district, southern Spain. This area is of particular interest as it is geologically and geometrically complex, ideal to demonstrate the importance of accurate corrections for spectral mapping.

With original sulfide reserves of 133 million tons, Tharsis is considered one of the most important mines in the Iberian Pyrite Belt and can be traced back to Chalcolithic and Roman times. In the SB open pit, both underground and opencast mining was conducted between 1866 and the 1990s [30], [31], and there is ongoing exploration work nearby. The main mineral commodity was pyrite, but the stockwork and gossan also contain As–Co–S, tellurides, and gold [31]. The massive sulfide ores form multiple thin but laterally extensive lenses embedded within a shale-dominated volcanic-sedimentary complex [31]. The mineralized sequence crops out discontinuously for more than 25 km along with the core of a regional E-W-trending antiform and can be subdivided into four south-dipping, thrust-bounded tectonic

TABLE I
ACQUISITION PARAMETERS FOR THE TWO HYPERSPECTRAL SCENES USED FOR THIS STUDY

Lighting	Location	Time	Exposure (ms)	Averaged scans	Sun azimuth	Sun elevation
Sunny	37.596512° N -7.120534° E	9/3/2020 16:10 UTC	VNIR: 21 SWIR: 20	3	241°	25°
Overcast	37.596501° N -7.120522° E	12/3/2020 10:22 UTC	VNIR: 92 SWIR: 52	3	133°	37°

a. Scene under sunny conditions
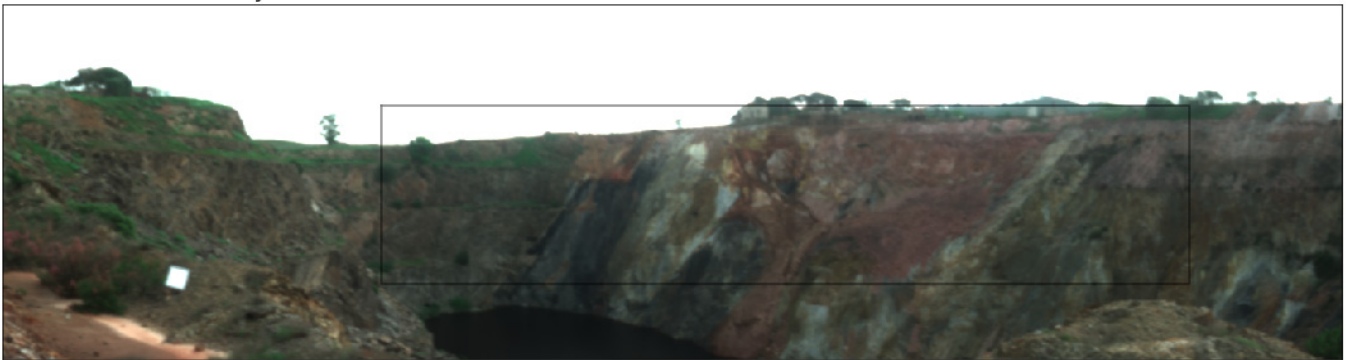

b. Scene under cloudy conditions

Fig. 2. True-color composites of hyperspectral radiance measured under (a) sunny conditions and (b) overcast conditions. Blue circles in (a) were used to characterize skylight for Fig. 5(a). Note that the scenes were captured from different imaging locations. The white rectangle shows the region that was projected (based on scene geometry and imaging locations) onto the cloudy scene (black rectangle) to allow pixelwise comparison in the following figures. This rectangle is 280 m across and 200–350 m from the viewing location. The Spectralon R90 calibration panels can be seen in (a) lower right and (b) lower left of each image.

units, comprising, from south to north: 1) slates and sandstones of the parautochthonous Phyllite-Quartzite (PQ) Group; 2) massive sulfides and slates; 3) slate and basalt with bodies of hydrothermal breccias; and 4) rhyodacite sills intruding slates. Both the shales and the felsic volcanic rocks adjacent to each ore body are affected by pervasive hydrothermal alteration, which manifests as ankeritization, sericitization, and silicification. The stockwork, located stratigraphically below the massive sulfides, is associated with the chloritic alteration.

Hyperspectral data covering the eastern side of the SB pit were acquired on two different days (see Fig. 2)—once in the afternoon during clear conditions, and once under bright but overcast conditions (see Table I). While overcast lighting is not ideal, requiring longer exposures and resulting in significantly more noise in the SWIR, it causes minimal topographic effects. Hence, the visible near-infrared (VNIR) part of this dataset provides an important baseline that is

almost free of topographic effects. We compare this baseline with corrected results from the sunny scene to evaluate the effectiveness of different topographic corrections.

Both scans were acquired using a tripod-mounted Asia FENIX (Specim, Oulu, Finland) hyperspectral pushbroom scanner. This camera has a wavelength range of 380–2500 nm at a spectral resolution of 3.5 nm in the VNIR and 12 nm in the SWIR. Its vertical field of view is 32.3° (over 384 pixels), and it has a maximum horizontal scanning angle of 130°. In this study, the camera was operated using a spectral binning model of 4/1 (for the VNIR/SWIR, respectively), resulting in 450 bands. Raw images were converted to radiance by subtracting dark-reference measurements and applying a sensor-specific correction to convert digital numbers to radiance. Finally, the signal-to-noise ratio for each scene was boosted by averaging the three individual scans captured sequentially during each acquisition.

a. Overcast scene after ELC
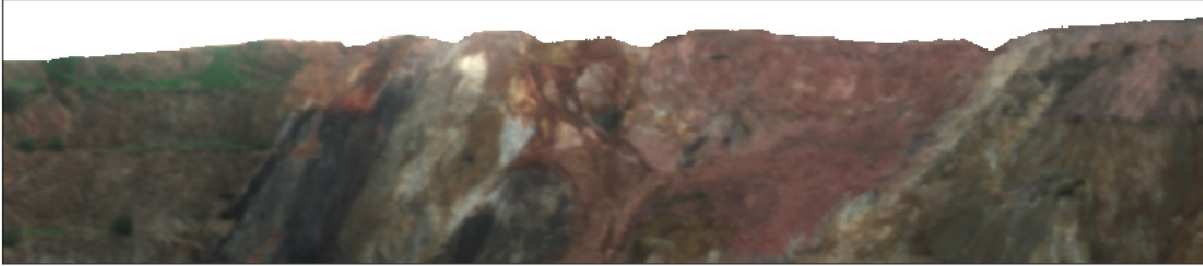
b. Sunny scene after ELC
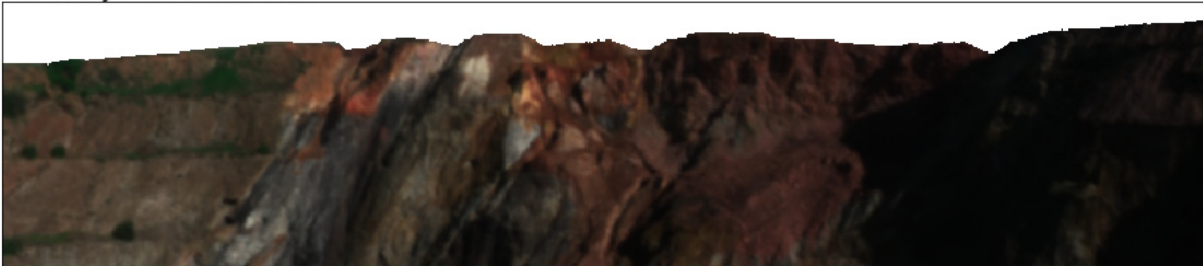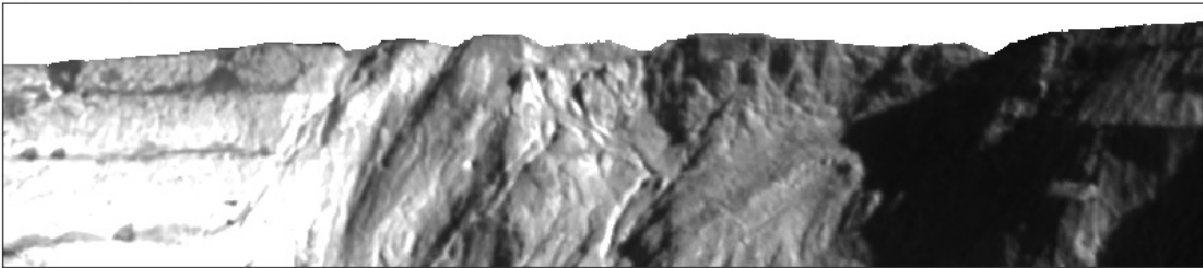
c. Illumination factor (for 680 nm)

Fig. 3. True-color composites of reflectance estimates derived using ELC from (a) overcast scene, which contains limited topographic effects, and (b) sunny scene, which contains significant topographic effects. The two can be compared to estimate this topographic effect (c), which, in this case, accounts for about half (55%) of the total variance in the sunny scene.

Two hundred and twenty-two geotagged ground images were also captured using a Nikon D850 (sensor size of 35.9 mm × 24.0 mm, with a pixel count of 7360 × 4912) and a Zeiss Milvus 2/35 ZF.2 lens, to construct a 3-cm resolution point cloud of the open pit using the structure from motion (SfM) workflow implemented in Agisoft Metashape v1.6. The resulting point cloud with relative positioning was subsequently georeferenced using a publicly available 2 m LiDAR dataset (PNOA-LiDAR, Instituto Geográfico Nacional, http://pnoa.ign.es) using the iterative closest point (ICP) algorithm integrated into CloudCompare v. 2.9.

The hyperspectral data were backprojected onto the photogrammetric point cloud using *hylite* [15]. *hylite* matches keypoints between each image and the point cloud using scale invariant feature transform (SIFT) [32] and then solves the Perspective-n-Point (PnP) problem using Open-CV [33] to accurately extract camera position and viewing direction. Points in the cloud were then projected onto each image to identify pixel correspondence, and the two datasets are aligned by projecting spectra from the sunny dataset onto the same pixel grid as the overcast dataset (see Fig. 2).

Finally, calibration spectra were extracted by averaging pixels for each image with a Spectralon R90 reference panel (made from a standardized polytetrafluoroethylene (PTFE) material with a stable reflectivity of 90% across the spectral range of interest) that was placed in the scene during each acquisition. The panels' orientations were estimated using *hylite*, based on their projected 2-D shape and known (square) geometry, and the associated sky view factor calculated by assuming the panel was relatively unoccluded and evaluating

$$a^{\text{panel}} = \mathbf{n_{panel}} \cdot \mathbf{s}. \tag{18}$$

A. Atmospheric (Only) Correction

It is standard practice with outcrop hyperspectral data to convert radiance to reflectance using an empirical line calibration (ELC; [20]). This uses calibration panel spectra to correct for the spectral composition of downwelling sunlight, which is (incorrectly) assumed to be distributed uniformly across the scene.

An ELC was applied to each of the scenes (see Fig. 3) to provide a reference with which to evaluate our novel correction algorithm. The difference between the independent reflectance estimates derived from the sunny and overcast scenes gives a useful approximation of the topographic effect that we aim
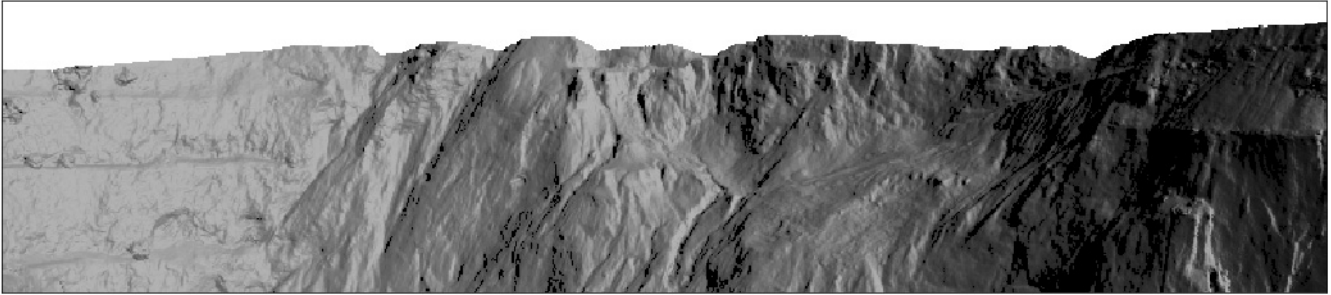
a. Illumination Type

Sun
Cast-shadow
Self-shadow
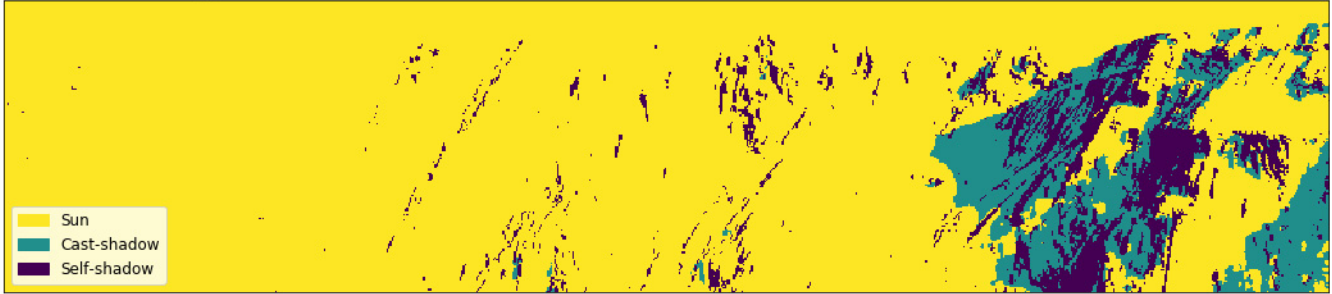
b. Lambertian shading

c. Oren-Nayar shading

d. Sky view factor

Fig. 4. (a) Types of illumination, with pixels classified as shade pixels if their normal points away from the light source and shadow pixels if they are oriented toward it but occluded (see text for details). (b) Lambertian shading highlights the geometric detail captured by the digital outcrop model but overestimates the topographic effect. Hence, we used (c) Oren–Nayar model. Finally, per-point sky view factors calculated using CloudCompare are shown in (d).

to correct [see Fig. 3(c)]. This topographic effect accounts for 55% of the variance in the sunny scene, highlighting the need for accurate correction approaches.

B. Joint Illumination Correction

Geometric data from the 3-D photogrammetric point cloud, including surface normal vectors, viewing direction, and sky view factors estimated using the PCV plugin [27], were projected onto the coregistered hyperspectral images using *hylite*. The position of the Sun was calculated using the *astral* python package [34] based on acquisition time and location (see Table I) for each acquisition.

The sunny scene was divided into three illumination classes: illuminated (sun), cast-shadow, and self-shadow [see Fig. 4(a)]. Pixels with normal vectors directed away from downwelling sunlight were considered to be self-shadowed. Cast-shadows were identified by averaging reflectance between 400 and 450 nm, dividing by the average from 550 to
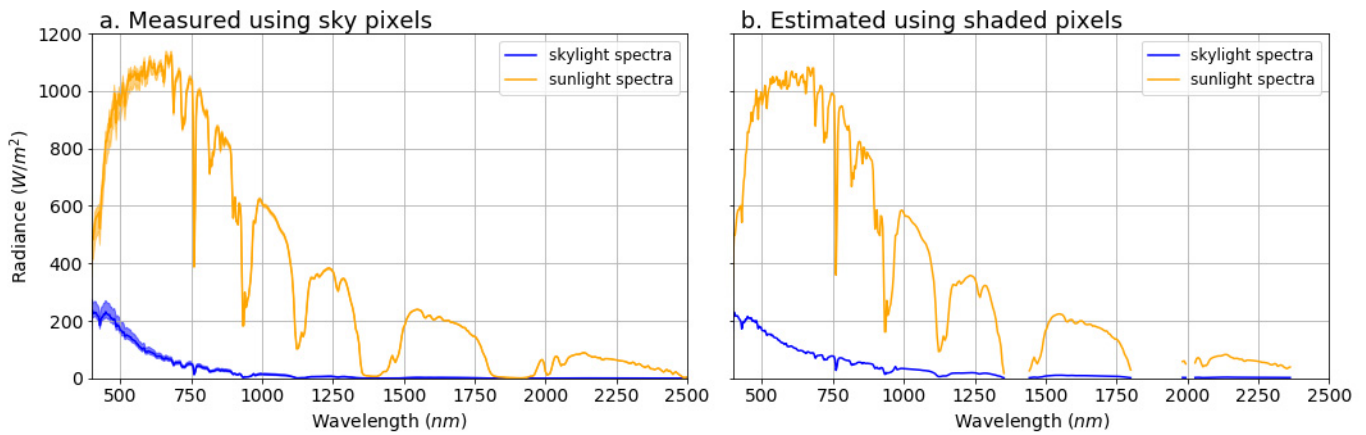
Fig. 5. Sunlight (yellow) and skylight (blue) spectra estimated by (a) sampling sky pixels and (b) using the estimation method proposed in Section II-B. The two spectra estimates are almost identical. The shaded area in (a) shows the 5th–95th percentile range of the sampled pixels.
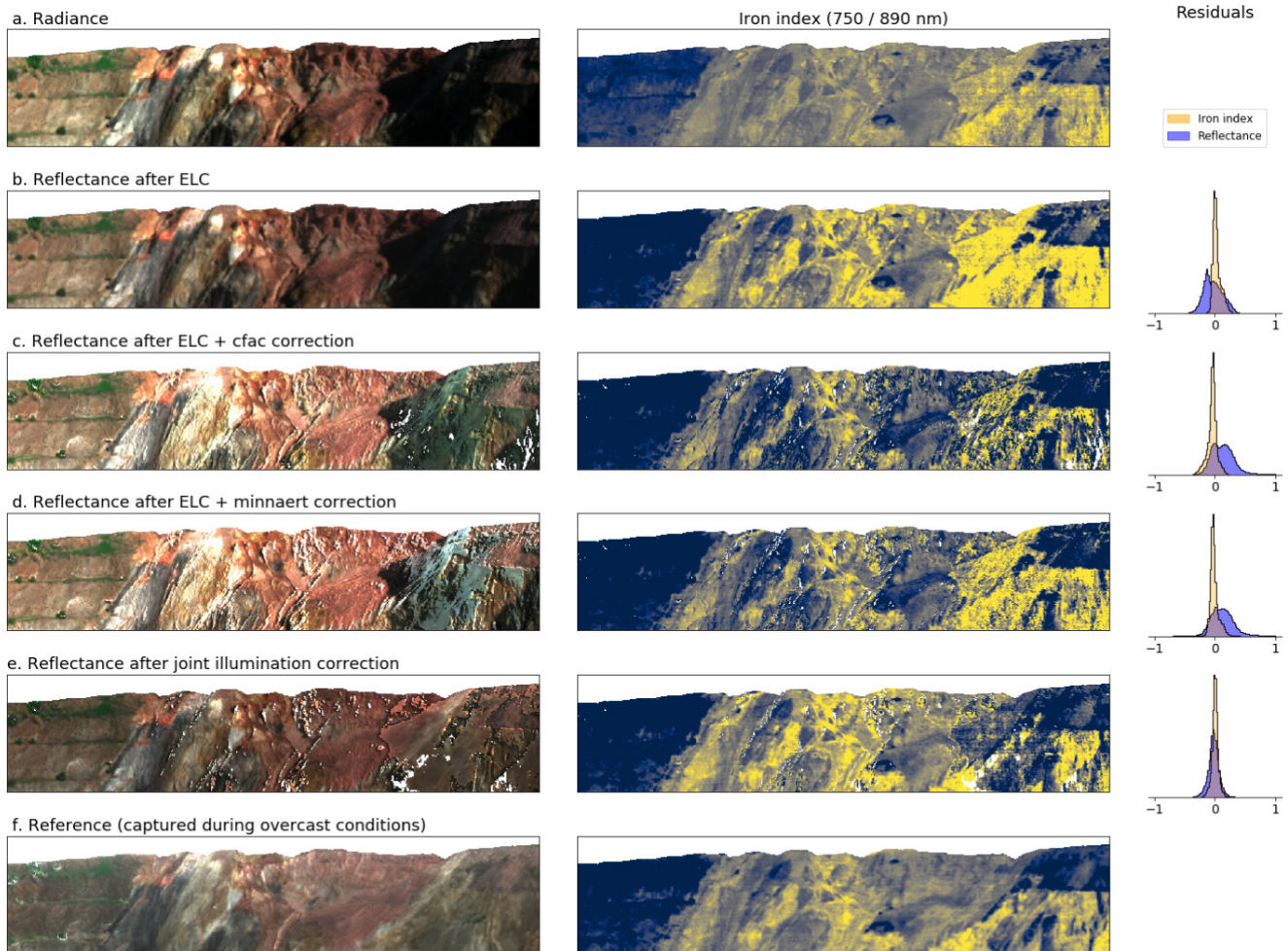

Fig. 6. Reflectance estimates derived by (a) correcting radiance measurements using standard procedures, including (b) only atmospheric correction (ELC), (c) c-factor correction, (d) minnaert correction, and (e) joint method proposed in this work. Reflectance estimates derived from (f) overcast scene, which has negligible topographic effects, were used to calculate residuals for each correction method. An iron index was also calculated, highlighting that inadequate topographic correction can cause significant errors in derivative products, such as band ratios. All RGB composites are visualized using a constant reflectance range of 0–0.6, while the iron indices are clipped from 1.0 to 1.25.

600 nm, and then applying a threshold such that values > 1.25 (and not already in the self-shadow class) were labeled as cast-shadow (cf. [29]).

Lambertian (1) and Oren–Nayar (2) reflection models were evaluated using the projected point normals, viewing direction, and Sun orientation [see Fig. 4(b) and (c)]. A large
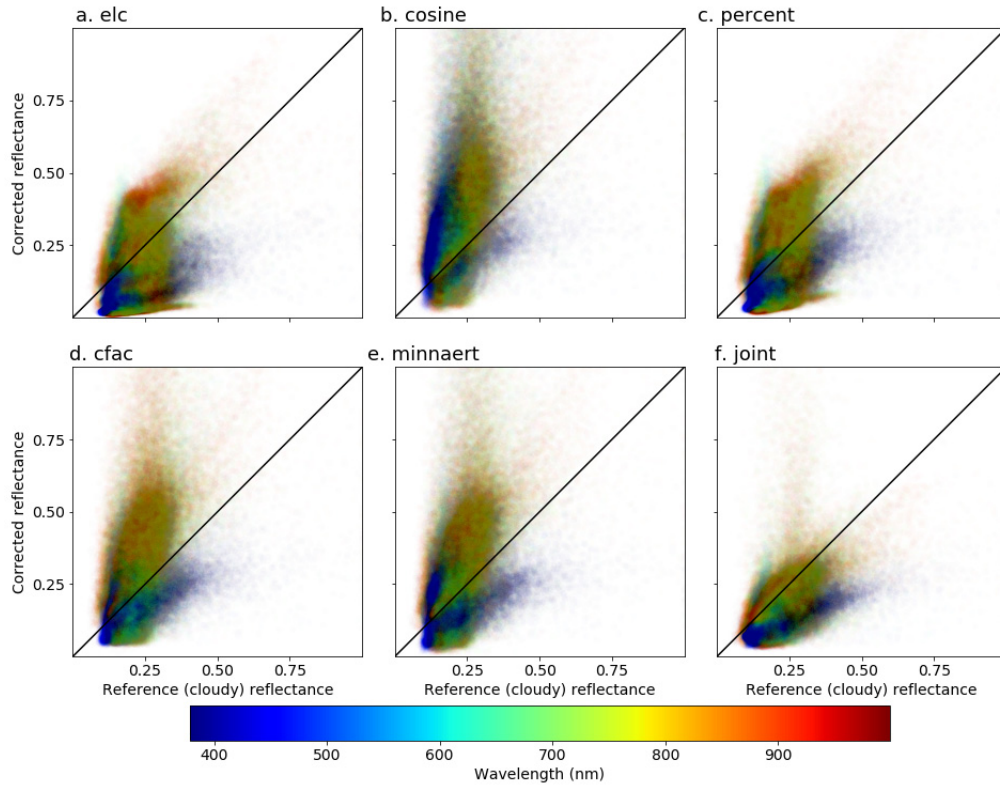
Fig. 7. Plots of reference (cloudy) reflectance versus results from the sunny dataset after applying (a) no topographic correction, and using (b) cosine, (c) percent, (d) c-factor, (e) minnaert, and (f) our joint correction method. The results highlight that: 1) there is a strong wavelength dependence in the correction results and 2) our joint correction method is closest to the reference data.

roughness ($\sigma = 40°$) was used in the Oren–Nayar model because the material in the scene is mostly coarse rubble or blocky, exfoliating rock formation. Comparing these results with Fig. 3(c), it is immediately evident that the Lambertian model overestimates the topographic effect, as has already been established [12]. Hence, the Oren–Nayar model was used for the rest of the correction steps.

The overcast scene was assumed to contain no direct sunlight, allowing the skylight spectra to be directly estimated from the calibration panel, and a joint illumination correction was applied using (8) to give baseline reflectance estimates. This correction is essentially equivalent to the ELC results [see Fig. 3(a)] but accounting for small illumination differences caused by variations in sky view factor.

Skylight spectra for the sunny scene were estimated using two independent methods: averaging the spectra of sky pixels in the original image [see Fig. 2(a)] and applying the estimation technique outlined in Section II-B. The results are almost identical (see Fig. 5), suggesting that (16) can be applied to real data with the inherent assumptions. This estimated skylight spectrum [see Fig. 5(b)] was used in (8) to derive corrected reflectance estimates.

Fig. 6 presents these corrected results, alongside those from uncorrected [see Fig. 6(a)], atmospherically corrected only [see Fig. 6(b)], and atmospheric (ELC) followed by state-of-the-art topographic corrections [see Fig. 6(c) and (d)]. Results from our joint illumination correction [see Fig. 6(e)] match the overcast reference [see Fig. 6(f)] best (see Table II).

TABLE II

ERROR ESTIMATES DERIVED BY COMPARING STATE-OF-THE-ART CORRECTION METHODS (SEE [12]) WITH THE REFLECTANCE BASELINE CAPTURED UNDER OVERCAST CONDITIONS. THESE SHOW THE MARKED IMPROVEMENT ACHIEVED BY OUR JOINT CORRECTION METHOD. INDEED, FOR THIS CASE, THE OTHER CORRECTION METHODS INCREASED ERROR RELATIVE TO A DATASET WITH NO TOPOGRAPHIC CORRECTION APPLIED (ELC ONLY) DUE TO SYSTEMATIC OVERCORRECTION THAT IS CAUSED ERRONEOUSLY HIGH REFLECTANCE ESTIMATES

Correction method	Median absolute percent error (%)	Median absolute error
ELC only	59.3	0.122
ELC + cosine	73.7	0.170
ELC + percent	55.3	0.118
ELC + c-factor	59.7	0.140
ELC + minnaert	59.9	0.137
Joint illumination correction	26.5	0.056

Indeed, the other evaluated topographic correction algorithms (cosine, percent, c-factor, and minnaert) all produce larger errors compared to an atmospheric-only ELC.

To highlight the spectral variations caused by the various corrections, an iron index (750/890 nm) was calculated. Although wavelength-independent topographic effects cancel out in this band-ratio, it is sensitive to changes in spectra slope and other distortions. The uncorrected datasets [see Fig. 6(a) and (b)] show systematic overestimation of iron content in shaded or shadowed regions, mostly notable on the right side
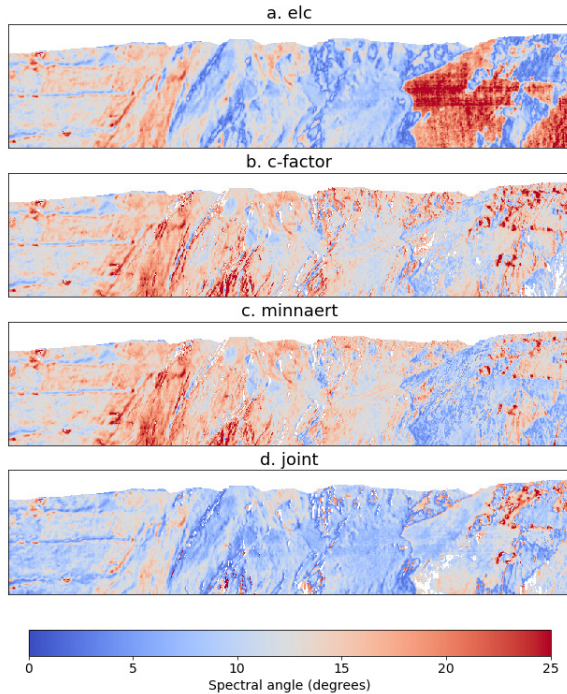
Fig. 8. Spectral angles between reference spectra and those derived using (a) ELC, (b) c-factor, (c) minnaert, and (d) our joint correction algorithm. Spectral angles are invariant to changes in bulk reflectance, so quantify differences in spectra shape.

of the scene. Topographic correction using the c-factor and minnaert methods gave results that also differ significantly from the reference dataset in these shadowed regions [see Fig. 6(c) and (d)]. Our joint correction gives results that closely match the reference [see Fig. 6(e)].

Plotting reference (cloudy) reflectance against corrected reflectance further highlights the spectral differences that result after applying different correction algorithms (see Fig. 7). Significantly, these differences are clearly wavelength dependent, with lower wavelengths generally falling below the unity line and higher wavelengths above it. This systematic spectral distortion was summarized by computing the spectral angle between each pixels' corrected reflectance spectra and the associated cloudy reference reflectance spectra (see Fig. 8). Our joint illumination correction produces results that most closely match the reference spectra (see Fig. 7) and show the least spectral distortion (see Fig. 8), though significant room for improvement remains.

Finally, the influence of these spectral differences on classification was explored using the spectral angle mapper (SAM) algorithm [35]. SAM has become widely applied by the remote sensing community (e.g., [36]–[38]) and, like the iron indices presented in Fig. 6, is insensitive to overall pixel brightness and so is generally assumed to be robust to topographic effects. The algorithm works by calculating the angles between pixel spectra (represented as a set of vectors) and reference spectra for each material. Reference spectra [see Fig. 9(a)] were calculated by choosing four 10×10 pixel squares for each lithology exposed at SB (volcanics, phyllite, massive sulphide, and gossan) [see Fig. 9(b)] and calculating their median spectra based on data from the reference (overcast)

image. The resulting classification [see Fig. 9(b)] was compared with results derived from topographically corrected data [see Fig. 9(c)–(f)]. While these classifications may not be geologically accurate (they have not been validated as ground-truth data are not available), they clearly demonstrate that the different correction methods influence classification results (see Table III), and thus, the SAM classifier is more sensitive to topographic effects than that has previously been assumed.

The classification without topographic correction [ELC only; see Fig. 9(c)] incorrectly labeled shadowed pixels as massive sulfide and only partially mapped the exposed ore body. Cosine and percent corrections do not alter spectra shape, so gave equivalent results (overall accuracy = 61%). Surprisingly, the c-factor [see Fig. 9(d)] and minnaert [see Fig. 9(e)] corrections degraded classification accuracy even further (to 55%) relative to the overcast dataset, and both methods completely failed to correctly identify massive sulfide. Classification results after our joint-correction method most closely matched the reference (68%) and correctly mapped the massive sulfide body.

IV. DISCUSSION

The results of the SB case study emphasize the need for appropriate illumination correction when using hyperspectral data to map surfaces with high spatial resolution devices. State-of-the-art corrections for satellite data (e.g., ATCOR; [19]) cannot be applied to the complex geometries and oblique viewing angles typical for digital outcrop models, so practitioners tend to either ignore topographic effects or to apply statistical corrections using e.g., the c-factor or minnaert algorithms (cf. [12]). Sophisticated codes, such as ATCOR, are proprietary, expensive, and not adapted.

Our results highlight the existence of significant errors when topographic effects are ignored. A naive interpretation of the iron index presented in Fig. 6 would incorrectly identify abundant iron oxides in shadowed regions of the SB pit. Similarly, the classification without topographic correction was 40% different from the reference classification, incorrectly labeling pixels in both well illuminated and shaded regions. C-factor and minnaert corrections did not improve classification accuracy (indeed, they degraded it by another 5%–10%). We suggest that this is because the standard corrections applied to convert radiance to reflectance (e.g., ELC) and established topographic corrections, such as c-factor and minnaert, do not properly account for the difference in illumination spectra between shaded and fully illuminated regions.

Despite its relative simplicity, the joint correction method that we propose successfully corrects these illumination differences by modeling the mixture of skylight and sunlight at each point in the scene. Ideally, downwelling skylight can be measured directly using a completely shaded calibration panel although our results show that the estimation method described in Section III-B can provide a reasonable approximation (e.g., for legacy data). The iron index calculated using this method closely matches the reference dataset (see Fig. 6), and the results could be further improved by better modeling the distribution of, e.g., partially shaded pixels in the penumbra
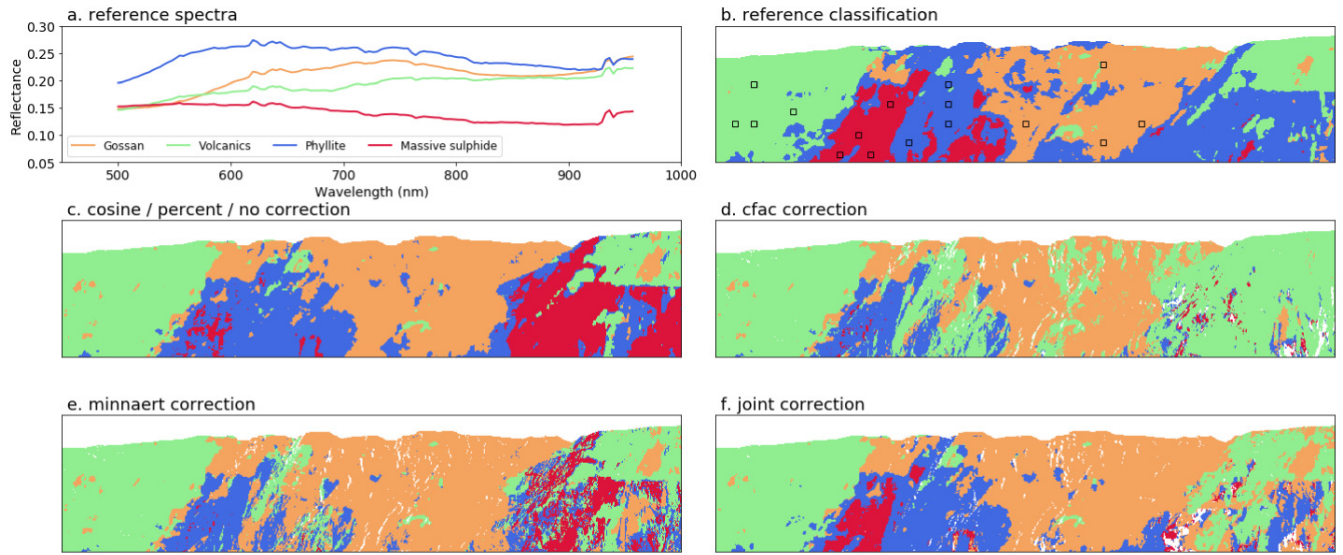
Fig. 9. Lithological classifications of the SB open pit using (a) SAM and reference spectra defined by calculating the median spectra of manually selected regions (black squares) in (b) reference dataset. Classification results from the sunny scene corrected using (c)–(f) various methods show significant variability, again emphasizing the importance of appropriate topographic correction. The classification from (f) our joint correction method best matches the reference classification (cf. Table III), which was calculated using the overcast dataset.

TABLE III

ESTIMATES OF THE CLASSIFICATION ERROR CAUSED BY TOPOGRAPHIC EFFECTS, CALCULATED BY COMPARING THE REFERENCE SCENE CLASSIFICATION WITH EQUIVALENT ONES DERIVED USING TOPOGRAPHICALLY CORRECTED DATA. RESULTS FOR ELC ONLY, COSINE, AND PERCENT CORRECTIONS ARE IDENTICAL AS THESE CORRECTIONS DO NOT INFLUENCE SPECTRA SHAPE, SO ARE INDISTINGUISHABLE USING SAM. C-FACTOR AND MINNAERT CORRECTIONS SIGNIFICANTLY DEGRADED CLASSIFICATION ACCURACY, WHILE OUR JOINT APPROACH IMPROVED IT BY ALMOST 10% RELATIVE TO NO TOPOGRAPHIC CORRECTION

Correction method	Overall accuracy (%)	Kappa score
ELC only	61	0.48
ELC + cosine	61	0.48
ELC + percent	61	0.48
ELC + c-factor	51	0.32
ELC + minnaert	56	0.4
Joint illumination correction	68	0.57

of cast shadows. Similarly, the classification results after joint correction showed a close match to the reference dataset, except in deeply shadowed regions where signal-to-noise ratios are presumably very low.

The error estimates presented in Tables II and III also highlight the improved performance of our correction method. While these must be treated cautiously, given uncertainties regarding the reference reflectances and associated classification from the cloudy dataset, they do highlight the tendency of established methods to overcorrect incidence angle effects. We suggest that this is caused by systematic underestimation of the reflected light fraction and associated overestimation of material reflectance. This could be explained by their failure to account for reflected skylight, and the resulting positive bias is evident in the residual histograms shown in Fig. 6. It has been well established that the cosine correction overcorrects illumination differences [12], [39], [40]; however, to the best of our

knowledge, this is the first time that a similar overcorrection has been demonstrated for the other approaches.

The estimated topographic effect shown in Fig. 3, which accounts for 55% of the variance in the uncorrected dataset, also suggests that more complex shading models are needed if we are to achieve accurate topographic corrections. The pit-face on the right-hand side of the scene [see Fig. 3(c)], which is almost perpendicular to both the viewing direction and illumination direction, appears to be more reflective than predicted by a Lambertian or Oren–Nayar model (see Fig. 4), presumably due to specular reflection. Adjacency effects caused by reflections from nearby outcrops and water filling the bottom of the open pit may also be important though, without a complex global illumination model, these are difficult to characterize.

The computer graphics community has developed a variety of global illumination models [41] that could be adapted for topographic correction, providing an interesting avenue for future research. The methods described in Section II for characterizing downwelling sun and skylight radiance spectra could be used to provide the necessary input for these and potentially lead to significantly better corrections (cf. Fig. 7). Empirical adjustments, such as the c-factor correction, could also be included to fit global illumination models to observed radiance, as could statistical methods for estimating path radiance (e.g., [42]).

In the meantime, the SB case study demonstrates that our proposed method provides a robust, practical correction that delivers usable results. A recent study [15] successfully employed the joint correction method presented here to map lithology and alteration minerals in an open-pit mine and emphasized the importance of this correction to achieving accurate results. A rapidly growing amount of hyperspectral data is becoming available to geoscientists, but the range of

sensors, acquisition conditions, and ground sampling distances (from laboratory to satellite) require accurate corrections [12] if data are to be quantitatively compared and synthesized.

To facilitate the use of our topographic correction method by the outcrop sensing community, we have added it (along with appropriate documentation and usage instructions) to the open-source *hylite* toolbox, a python package for correcting and analyzing hyperspectral datasets [15]. We suggest that the characterization of downwelling skylight using a completely shaded calibration panel is worthwhile and should become standard practice. Multiexposure high dynamic range techniques could also be employed to maximize signal-to-noise ratios across the whole scene when illumination contrasts are large.

V. Conclusion

In conclusion, outcrop geometry causes important illumination differences that, if uncorrected, hinder the identification and mapping of outcrop mineralogy. By separating downwelling sun and skylight spectra and modeling their contribution to each pixel's illumination, our novel and open-source correction method achieves substantially better results than established ones. Absolute reflectance values are closer to reference measurements, and spectral distortions caused by different mixtures of sun and skylight are corrected for. Hence, we hope that it is a valuable tool for the outcrop remote sensing community. These corrections are fundamental for any proximal hyperspectral acquisition, especially those performed by drones.

Acknowledgment

References

[1] D. Hodgetts, "Laser scanning and digital outcrop geology in the petroleum industry: A review," *Mar. Petroleum Geol.*, vol. 46, pp. 335–354, Sep. 2013. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0264817213000494

[2] S. P. Bemis *et al.*, "Ground-based and UAV-based photogrammetry: A multi-scale, high-resolution mapping tool for structural geology and paleoseismology," *J. Struct. Geol.*, vol. 69, pp. 163–178, Dec. 2014. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0191814114002429

[3] P. R. Nesbit, A. Boulding, C. Hugenholtz, P. Durkin, and S. Hubbard, "Visualization and sharing of 3D digital outcrop models to promote open science," *GSA Today*, vol. 30, no. 6, pp. 4–10, Jun. 2020. [Online]. Available: https://www.geosociety.org/GSA/Publications/GSA_Today/GSA/GSAToday/scie%nce/G425A/article.aspx

[4] S. T. Thiele, L. Grose, A. Samsu, S. Micklethwaite, S. A. Vollgger, and A. R. Cruden, "Rapid, semi-automatic fracture and contact mapping for point clouds, images and geophysical data," *Solid Earth*, vol. 8, no. 6, p. 1241, 2017.

[5] T. J. B. Dewez, D. Girardeau-Montaut, C. Allanic, and J. Rohmer, "Facets: A cloudcompare plugin to extract geological planes from unstructured 3D point clouds," *Int. Arch. Photogramm., Remote Sens. Spatial Inf. Sci.*, vol. XLI-B5, pp. 799–804, Jun. 2016.

[6] P. R. Nesbit, P. R. Durkin, C. H. Hugenholtz, S. M. Hubbard, and M. Kucharczyk, "3-D stratigraphic mapping using a digital outcrop model derived from UAV images and structure-from-motion photogrammetry," *Geosphere*, vol. 14, no. 6, pp. 2469–2486, Oct. 2018. [Online]. Available: https://pubs.geoscienceworld.org/gsa/geosphere/article/14/6/2469/559082%/3-D-stratigraphic-mapping-using-a-digital-outcrop

[7] T. H. Kurz, S. J. Buckley, J. A. Howell, and D. Schneider, "Integration of panoramic hyperspectral imaging with terrestrial lidar data," *Photogramm. Rec.*, vol. 26, no. 134, pp. 212–228, Jun. 2011, doi: 10.1111/j.1477-9730.2011.00632.x.

[8] R. J. Murphy, Z. Taylor, S. Schneider, and J. Nieto, "Mapping clay minerals in an open-pit mine using hyperspectral and LiDAR data," *Eur. J. Remote Sens.*, vol. 48, no. 1, pp. 511–526, Jan. 2015.

[9] S. J. Buckley *et al.*, "LIME: Software for 3-D visualization, interpretation, and communication of virtual geoscience models," *Geosphere*, vol. 15, no. 1, pp. 222–235, Jan. 2019, doi: 10.1130/GES02002.1.

[10] G. M. Dering, S. Micklethwaite, S. T. Thiele, S. A. Vollgger, and A. R. Cruden, "Review of drones, photogrammetry and emerging sensor technology for the study of dykes: Best practises and future potential," *J. Volcanology Geothermal Res.*, vol. 373, pp. 148–166, Mar. 2019. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0377027318302294

[11] S. J. Buckley, T. H. Kurz, J. A. Howell, and D. Schneider, "Terrestrial lidar and hyperspectral data fusion products for geological outcrop analysis," *Comput. Geosci.*, vol. 54, pp. 249–258, Apr. 2013. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0098300413000320

[12] S. Jakob, R. Zimmermann, and R. Gloaguen, "The need for accurate geometric and radiometric corrections of drone-borne hyperspectral data for mineral exploration: MEPHySTo—A toolbox for pre-processing drone-borne hyperspectral data," *Remote Sens.*, vol. 9, no. 1, p. 88, Jan. 2017. [Online]. Available: https://www.mdpi.com/2072-4292/9/1/88

[13] M. Kirsch *et al.*, "Integration of terrestrial and drone-borne hyperspectral and photogrammetric sensing methods for exploration mapping and mining monitoring," *Remote Sens.*, vol. 10, no. 9, p. 1366, Aug. 2018. [Online]. Available: https://www.mdpi.com/2072-4292/10/9/1366

[14] M. Kirsch *et al.*, "Hyperspectral outcrop models for palaeoseismic studies," *Photogramm. Rec.*, vol. 34, no. 168, pp. 385–407, Dec. 2019. [Online]. Available: https://onlinelibrary.wiley.com/doi/pdf/10.1111/phor.12300

[15] S. T. Thiele *et al.*, "Multi-scale, multi-sensor data integration for automated 3-D geological mapping," *Ore Geol. Rev.*, vol. 136, Sep. 2021, Art. no. 104252.

[16] M. J. Duggin, "Review article factors limiting the discrimination and quantification of terrestrial features using remotely sensed radiance," *Int. J. Remote Sens.*, vol. 6, no. 1, pp. 3–27, Jan. 1985, doi: 10.1080/01431168508948420.

[17] Y. Kawata, S. Ueno, and T. Kusaka, "Radiometric correction for atmospheric and topographic effects on landsat MSS images," *Int. J. Remote Sens.*, vol. 9, no. 4, pp. 729–748, Apr. 1988, doi: 10.1080/01431168808954889.

[18] C. Proy, D. Tanre, and P. Y. Deschamps, "Evaluation of topographic effects in remotely sensed data," *Remote Sens. Environ.*, vol. 30, no. 1, pp. 21–32, 1989. [Online]. Available: https://www.sciencedirect.com/science/article/pii/0034425789900448

[19] R. Richter and D. Schläpfer, "Geo-atmospheric processing of airborne imaging spectrometry data. Part 2: Atmospheric/topographic correction," *Int. J. Remote Sens.*, vol. 23, no. 13, pp. 2631–2649, 2002, doi: 10.1080/01431160110115834.

[20] G. M. Smith and E. J. Milton, "The use of the empirical line method to calibrate remotely sensed data to reflectance," *Int. J. Remote Sens.*, vol. 20, no. 13, pp. 2653–2662, 1999, doi: 10.1080/014311699211994.

[21] C. Wang and S. W. Myint, "A simplified empirical line method of radiometric calibration for small unmanned aircraft systems-based remote sensing," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 8, no. 5, pp. 1876–1885, May 2015.

[22] M. Oren and S. K. Nayar, "Seeing beyond Lambert's law," in *Computer Vision—ECCV* (Lecture Notes in Computer Science), J.-O. Eklundh, Ed. Berlin, Germany: Springer, 1994, pp. 269–280.

[23] D. Carrea, A. Abellan, F. Humair, B. Matasci, M.-H. Derron, and M. Jaboyedoff, "Correction of terrestrial LiDAR intensity channel using Oren–Nayar reflectance model: An application to lithological differentiation," *ISPRS J. Photogramm. Remote Sens.*, vol. 113, pp. 17–29, Mar. 2016. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0924271615002658

[24] M. Tarini, P. Cignoni, and R. Scopigno, "Visibility based methods and assessment for detail-recovery," in *Proc. IEEE Trans. Ultrason., Ferroelectr., Freq. Control*, 2003, pp. 457–464.

[25] P. P. Kastendeuch, "A method to estimate sky view factors from digital elevation models," *Int. J. Climatol.*, vol. 33, no. 6, pp. 1574–1578, May 2013, doi: 10.1002/joc.3523.

[26] S. M. An *et al.*, "Three-dimensional point cloud based sky view factor analysis in complex urban settings," *Int. J. Climatol.*, vol. 34, no. 8, pp. 2685–2701, Jun. 2014, doi: 10.1002/joc.3868.

[27] M. A. Berthaume, J. Winchester, and K. Kupczik, "Ambient occlusion and PCV (portion de ciel visible): A new dental topographic metric and proxy of morphological wear resistance," *PLoS ONE*, vol. 14, no. 5, May 2019, Art. no. e0215436. [Online]. Available: https://journals.plos.org/plosone/article?id=10.1371/journal.pone.02154%36

[28] O. Friman, G. Tolt, and J. Ahlberg, "Illumination and shadow compensation of hyperspectral images using a digital surface model and non-linear least squares estimation," *Proc. SPIE*, vol. 8180, Oct. 2011, Art. no. 81800Q. [Online]. Available: https://www.spiedigitallibrary.org/conference-proceedings-of-spie/8180/%81800Q/Illumination-and-shadow-compensation-of-hyperspectral-images-using-a-di%gital/10.1117/12.898084.short

[29] A. Sanin, C. Sanderson, and B. C. Lovell, "Shadow detection: A survey and comparative evaluation of recent methods," *Pattern Recognit.*, vol. 45, no. 4, pp. 1684–1695, Apr. 2012. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0031320311004043

[30] R. M. González, M. Olías, F. Macías, C. R. Cánovas, and R. F. de Villarán, "Hydrological characterization and prediction of flood levels of acidic pit lakes in the Tharsis mines, Iberian Pyrite Belt," *J. Hydrol.*, vol. 566, pp. 807–817, Nov. 2018. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0022169418307406

[31] F. Tornos, M. Solomon, C. Conde, and B. F. Spiro, "Formation of the Tharsis massive sulfide deposit, Iberian Pyrite Belt: Geological, lithogeochemical, and stable isotope evidence for deposition in a brine pool," *Econ. Geol.*, vol. 103, no. 1, pp. 185–214, Jan. 2008. [Online]. Available: https://pubs.geoscienceworld.org/segweb/economicgeology/article/103/1/1%85/127950/Formation-of-the-Tharsis-Massive-Sulfide-Deposit

[32] D. G. Lowe, "Object recognition from local scale-invariant features," in *Proc. IEEE Int. Conf. Comput. Vis.*, vol. 2. Sep. 1999, pp. 1150–1157.

[33] G. Bradski and A. Kaehler, *Learning OpenCV: Computer Vision With the OpenCV Library*. Sebastopol, CA, USA: O'Reilly Media, 2008.

[34] S. Kennedy. (Mar. 2021). *Sffjunkie/Astral*. [Online]. Available: https://github.com/sffjunkie/astral

[35] F. A. Kruse *et al.*, "The spectral image processing system (SIPS)-interactive visualization and analysis of imaging spectrometer data," *Remote Sens. Environ.*, vol. 44, nos. 2–3, pp. 145–163, 1993. [Online]. Available: https://www.sciencedirect.com/science/article/pii/003442579390013N

[36] X. Chen, T. A. Warner, and D. J. Campagna, "Integrating visible, near-infrared and short-wave infrared hyperspectral and multispectral thermal imagery for geological mapping at cuprite, Nevada," *Remote Sens. Environ.*, vol. 110, no. 3, pp. 344–356, Oct. 2007. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0034425707001228

[37] C. Hecker, M. van der Meijde, H. van der Werff, and F. D. van der Meer, "Assessing the influence of reference spectra on synthetic SAM classification results," *IEEE Trans. Geosci. Remote Sens.*, vol. 46, no. 12, pp. 4162–4172, Dec. 2008.

[38] X. Zhang and P. Li, "Lithological mapping from hyperspectral data by improved use of spectral angle mapper," *Int. J. Appl. Earth Observ. Geoinf.*, vol. 31, pp. 95–109, Sep. 2014. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0303243414000579

[39] A. Twele and S. Erasmi, "Evaluating topographic correction algorithms for improved land cover discrimination in mountainous areas of central Sulawesi," *Remote Sens. GIS Environ. Stud.*, vol. 113, pp. 287–295, Jan. 2005.

[40] I. Sola, M. Gonzalez-Audicana, J. Alvarez-Mozos, and J. L. Torres, "Synthetic images for evaluating topographic correction algorithms," *IEEE Trans. Geosci. Remote Sens.*, vol. 52, no. 3, pp. 1799–1810, Mar. 2014.

[41] T. Ritschel, C. Dachsbacher, T. Grosch, and J. Kautz, "The state of the art in interactive global illumination," *Comput. Graph. Forum*, vol. 31, no. 1, pp. 160–188, Feb. 2012, doi: 10.1111/j.1467-8659.2012.02093.x.

[42] S. Lorenz *et al.*, "Radiometric correction and 3D integration of long-range ground-based hyperspectral imagery for mineral exploration of vertical outcrops," *Remote Sens.*, vol. 10, no. 2, p. 176, Jan. 2018. [Online]. Available: https://www.mdpi.com/2072-4292/10/2/176