Investigating impacts of calibration methodology and irradiance variations on lightweight drone-based sensor derived surface reflectance products

Dominic Fawcett, Karen Anderson
Investigating impacts of calibration methodology and irradiance variations on lightweight drone-based sensor derived surface reflectance products

Dominic Fawcett\textsuperscript{a,*}, Karen Anderson\textsuperscript{a}

\textsuperscript{a}Environment and Sustainability Institute, University of Exeter, Penryn, Cornwall, TR10 9FE, UK

\* fawcettdominic@gmail.com

ABSTRACT

The miniaturisation of multispectral sensors in recent years has resulted in a proliferation of applications particularly in vegetation-focused studies using lightweight drones. Multi-camera arrays (MCAs), capable of capturing information over different wavelength intervals using separate cameras with specific band-pass filters, are now commonplace in this field. However, data from MCAs require a considerable amount of geometric and radiometric corrections if high quality reflectance products are to be delivered. Some aspects of this workflow can be handled by commercial software packages (e.g. Pix4D and Agisoft Metashape), using black box algorithms, however radiometric uncertainties within products are not reported to the end-user by the software. We present the results of two experiments using a low-cost MCA complete with irradiance sensor (Parrot Sequoia), which set out to assess the accuracy and consistency of hemispherical-conical surface reflectance factors from MCA data. Using reference panels in the field, we found that the empirical line method (ELM) generated the smallest RMSEs (0.0037) when compared to simplified single-panel based workflows; while for the latter there was little difference between using a calibrated Spectralon\textsuperscript{®} panel or grey card imaged prior to the flight (0.0215 vs 0.0154 average over the four bands). Errors for a vegetated target within the survey flight were larger and comparable for all cases. Furthermore, a study on median vegetation index values for single vegetation canopies showed that illumination correction using irradiance data still yields significant differences in resulting values between two acquisitions during changing direct and diffuse irradiance conditions. We therefore highlight the importance of critical assessment prior to integrating drone derived MCA-measured reflectance factors into further geospatial workflows.

1 INTRODUCTION

The multi-camera array (MCA) solution to multi-spectral imaging is attractive due to the relatively low-cost and light weight of the required hardware, meaning the cameras can be deployed on small multi-rotor drones which offer unique flexibility. Furthermore, the fully integrated systems usually weigh in at \(<3\) kg making them exempt from regulations that control the deployments of heavier drone aircraft; and are therefore deployable by scientists at a large range of study sites. In particular the Parrot Sequoia camera has been applied in a number of recent vegetation related studies\textsuperscript{1,2,3}. However, in part due to fully integrated processing in commercial software (i.e. Pix4D), reports regarding the uncertainties within surface reflectance and derived indices are often lacking. In particular, the quality of the reference panel and its calibrated reflectance used to convert the generated pseudo-radiance orthomosaics to surface reflectance factors has been shown to...
be a potential large source of error due to degradation, an issue which is commonly encountered for spectroradiometric studies in the field \cite{4,5,6}.

Furthermore, important metadata such as temporal/spatial irradiance variations during acquisitions are not commonly reported, and instead there is often a reliance on the integrated irradiance compensation to provide homogenous products, although it has been previously observed that this does not always lead to consistent improvements\cite{7}.

Within this study, we assess two key constraints regarding the reliability of drone-mounted MCA derived surface reflectance factors:

1. the influence of the type of reflectance calibration method used; and
2. the compensation of irradiance variations.

Experiment 1 tested the hypothesis that the two most commonly employed approaches for the retrieval of surface reflectance from MCA image data would deliver the same information. First, a standard empirical line method (ELM) \cite{8} was used, using reference points of multiple panels of known reflectance imaged in-flight. Second, a simplified ELM (e.g. \cite{9,10}) implemented in software packages which relies on a single ground-based calibration image of a reference panel was used as a comparison. Panels for the in-flight ELM correction were self-manufactured while for the simplified ELM a Kodak grey card middle grey reference as well as a Spectralon® calibrated reflectance panel was used. Derived surface reflectance factors were compared to those of validation targets measured in-situ using an OceanOptics FLAME spectrometer measuring over the same spectral range as the MCA instrument.

The second experiment tested the hypothesis that in-flight irradiance information could eliminate variations stemming from changing illumination within the generated surface reflectance product. The impact of correcting for illumination variations based on irradiance sensor data was assessed by comparing surface reflectance factors and derived vegetation index values (normalised difference vegetation index (NDVI) and chlorophyll index (CHL)) from acquisitions with changing irradiance conditions due to patchy cloud cover.

2 MATERIALS AND METHODS

Both experiments were conducted using a low-cost MCA complete with irradiance sensor (Parrot Sequoia, Parrot, France) mounted on a 3DR Solo quadcopter using custom 3D printed mounts. The Sequoia is mounted at 3° off-nadir in forward direction to compensate the average forward tilt of the drone when flying at 5 m/s.

The Sequoia multispectral sensor consists of four cameras with different band-pass filters located in the green, red, red-edge and NIR regions of the spectrum with centre wavelengths at 550, 640, 735 and 790 nm and varying bandwidths. The approximated sensor response per band is visualised in Figure 1 along with a simulated reflectance spectrum representing healthy vegetation.

\footnote{Spectralon is a registered trademark of Labsphere, Inc.}
2.1 Experiment 1

2.1.1 Study site and drone flights

Data were acquired over a pasture and mixed woodland study area in South West UK during cloud-free conditions at 11:30 AM BST (UTC+1) on 4th July 2019. The drone was flown in a pre-programmed north-south oriented lawnmower pattern at 70 m elevation above ground at a speed of 5 m/s. This resulted in ~80% lateral and ~90% frontal overlap between images and a nominal ground resolution of 6.6 cm. Calibration panels deployed in the field were imaged at a lower elevation of 25 m resulting in a nominal ground resolution of 2.36 cm.

2.1.2 Calibration panels

A calibrated Spectralon panel of 42.5% average reflectance over the Sequoia wavelength range (SRT-40-050) as well as a calibrated grey card of ~20% average reflectance were imaged prior to the flight at three different exposures with the Sequoia Radiometric Calibration setting. For each band, the non-saturated image with the longest exposure was selected as reference.

Five reference panels of varying reflectances ranging from 2% to 42% (Figure 2) were levelled and imaged at a lower flying altitude of 25 m in order to resolve sufficient pure pixels per panel and reduce adjacency effects. Panels (c), (d) and (e) are plywood painted with Rust-Oleum® matte furniture paint while (f) and (g) are fabric panels manufactured by MosaicMill (Finland). All panels are designed to minimise bi-directional effects, though the actual BRDF was not assessed here.
2.1.3 Field spectral measurements

The reflectance of the five reference panels as well as one grass and one pseudo-invariant concrete surface were measured immediately prior to the drone acquisition (11:20-11:25 UTC+1), using an OceanOptics FLAME spectrometer and a calibrated Spectralon white reference panel (SRS-99-020, 99% reflectance across the Sequoia wavelength range). Measured reflectance spectra were post-processed by interpolating artefacts stemming from absorption features, the most prominent in the O₂A band (760 nm), and smoothing the spectra using Savitzky-Golay filtering \(^1\). In order to compare the reference values to the measured Sequoia band values, the spectra were convolved with each band’s spectral response function (SRF). Due to the true SRFs being unavailable, they were derived from filter transmission values per wavelength and the approximated camera CMOS sensor response (see Figure 1).

2.1.4 Drone derived HCRF

The following formulas for the calibration of Sequoia images using information recorded in image EXIF tags are based on information provided by the manufacturer\(^1\) and are fully integrated within the Pix4D software (with the exception of the standard ELM based on multiple panels). However, the implementation of the full processing pipeline remains black-box and cannot be detailed here. Lens distortion and vignetting corrections are also applied within the software (see Pix4D, 2019).

The digital numbers \(P\) as recorded per pixel by the Sequoia sensor are converted to radiance-equivalent values, here termed pseudo-radiance \(R\), following eq. (1) in Table 1. Here \(f\) is the focal ratio, \(\gamma\) the ISO and \(\varepsilon\) the exposure time in seconds whereas \(A\), \(B\) and \(C\) are camera specific calibration factors provided by the manufacturer.

For the single-panel simplified ELM calibration, \(R\) is converted to surface reflectance \(\rho\) using calibration coefficient \(K\) (eq. 3, Table 1). \(K\) is derived from the calibrated reflectance value \(\rho_{\text{ref}}\) of a reference panel imaged prior to the flight, as well as the average \(R\) over the surface of the imaged panel, termed \(\bar{R}_{\text{ref}}\) (eq. 2, Table 1).

Alternatively, the irradiance counts in DNs (\(C\)) recorded by the irradiance sensor are converted to pseudo-irradiance (\(R_{\text{irr}}\)) using eq. (4), where \(g\) is the gain and \(\varepsilon_{\text{irr}}\) is the exposure time of the irradiance sensor. \(R_{\text{irr}}\) can then be used to account for the variation in irradiance in respect to the single panel reference image and derive a \(\rho\) which is less susceptible to

![Figure 2: The grey card (a) and Spectralon panel on tripod (b) imaged pre-flight, and the five larger reference panels (c-g, with averaged reflectances of 0.2, 0.07, 0.39, 0.02 and 0.42 respectively) imaged at 25 m altitude, displayed in the NIR band.](image-url)
irradiance differences, using eq. (6) in Table 1. $K'$ is derived similarly to $K$ but includes the irradiance measured during the capture of the single panel calibration image $R_{\text{ref irr}}$ (eq. 5, Table 1). This factor also accounts for the differences in sensitivities and solid angles between the irradiance sensor and camera pixel measurements.

Table 1: Formulas used to convert digital numbers as measured by the Sequoia camera and irradiance sensor to surface reflectance. Left column: reflectance calibration not considering irradiance, Right column: Reflectance calibration with irradiance compensation.

<table>
<thead>
<tr>
<th>Pseudo-Radiance ($R$) from DNs ($P$):</th>
<th>Pseudo-Irradiance ($R_{\text{irr}}$) from DNs ($C$):</th>
</tr>
</thead>
<tbody>
<tr>
<td>$R = f^2 \frac{P-B}{\Delta^2+\varepsilon}$ (eq. 1)</td>
<td>$R_{\text{irr}} = \frac{C}{g^2_{\text{irr}}}$ (eq. 4)</td>
</tr>
<tr>
<td>Calibration coefficient using reference panel (simplified ELM):</td>
<td>Calibration coefficient using reference panel and measured irradiance (simplified ELM):</td>
</tr>
<tr>
<td>$K = \frac{\rho_{\text{ref}}}{R_{\text{ref}}}$ (eq. 2)</td>
<td>$K' = \rho_{\text{ref}} \frac{R_{\text{ref irr}}}{R_{\text{ref}}}$ (eq. 5)</td>
</tr>
<tr>
<td>$\rho = KR$ (eq. 3)</td>
<td>$\rho = K' \frac{R}{R_{\text{irr}}}$ (eq. 6)</td>
</tr>
</tbody>
</table>

For the ELM calibration method, surface reflectance $\rho$ was predicted from pseudo radiance $R$ using a linear model based on a dark and bright reference target. The bright targets used were different for the visible bands due to saturation (here 2% and 25% reflective panels for the Green and Red bands, 2% and 42% reflective panels for the red edge and NIR bands).

2.2 Experiment 2

2.2.1 Study site and drone flights

These data were acquired over a 10 year old oil-palm plantation in Sarawak, Malaysian Borneo between 10:30 and 11:20 AM UTC+8 on 3rd February 2018. Two identical flights were conducted following a lawnmower pattern flight plan at 100 m elevation above ground and 5 m/s speed with image acquisitions triggered by intervalometer which ensured 80% lateral and >90% frontal image overlap and yielded 9.42 cm ground resolution. For the first flight, here referred to as F1, illumination conditions were more uniform throughout the flight than compared to the second flight (F2). However, while for F1 irradiance was mostly diffuse, both flights showed variations in irradiance.

2.2.2 Calibration panels

The flights were calibrated using a Spectralon white reference panel (SRT-99-050, 99% reflectance over the Sequoia wavelength regions) imaged during overcast conditions at three different exposures with the Sequoia Radiometric Calibration setting. For each band, the non-saturated image with the longest exposure was selected as reference.

2.2.3 Drone derived HCRF

The simplified ELM radiometric calibration is the same as described for Experiment 1.
The pseudo radiance $R$ from both flights was calibrated to surface reflectance by alternatively using equations (3) and (6), resulting in two surface reflectance products each, one which considers measured irradiance and one which uses only the Spectralon (SRT-99-050, 99% reflective) panel as reference.

### 2.2.4 Per-palm NDVI and CHL values

The NDVI and CHL indices were calculated for both acquisitions as follows:

\[
NDVI = \frac{NIR - Red}{NIR + Red} \quad \text{(eq. 7)} \quad CHL = \frac{NIR}{Red \, edge} - 1 \quad \text{(eq. 8)}
\]

Eq. (7) is based on the work by Gitelson et al. (2006).

The locations of 496 palm tree tops were derived using structure-from-motion photogrammetry in a previous study conducted on this site. Median NDVI and CHL values were extracted within a 2.5 m buffer around each point, the buffer size was chosen to minimise the impact of overlapping palm fronds or visible ground between palms.

### 3 RESULTS

#### 3.1 Experiment 1

The measured surface reflectance averaged over the area of each target is plotted against the field-measured reference reflectance (HCRF) in Figure 3, while corresponding root mean square errors and standard deviations over the reference panel areas are recorded in Table 2.

The two brightest targets with >30% average reflectance saturated in the green and red band and were therefore omitted from this analysis. This represents a limitation of the minimum automatic exposure time of the Parrot Sequoia sensor, which is designed to capture small variations within vegetation reflectance which is very low in the green and red bands due to absorption. The dark and bright targets used to construct the empirical line for the ELM were also omitted from the further analysis.

As is apparent from Table 2, for the reference panels the mean RMSEs over all bands is considerably lower for the ELM calibration compared to the single-panel approaches. Between the two panels used for the simplified ELM, the Kodak grey card performs better than the Spectralon (SRT-40-050) panel with 0.0154 mean RMSE over all bands compared to 0.0215 for the Spectralon panel. In the red edge and NIR bands, larger deviations for the single panel calibrations can be observed at lower reflectance. As these panels were imaged at lower altitude to prevent adjacency effects and the influence of water vapour and aerosols is assumed to be negligible at these altitudes, one possible source of error is dark-current which may not be adequately accounted for within the calibration workflow. This does not appear to influence the visible bands.

Errors are highest for the grass meadow target in the red edge and NIR bands, with mean absolute errors up to 6.7% in the red edge when using the grey card. These errors can have a substantial impact, particularly for the derivation of chlorophyll content from indices. The source of this anomaly is unclear and may be related to non-representative measurements in the field or uncertainties within the sensor response used to convolve the measured spectra to representative Sequoia measurements. The latter should be investigated as it has far reaching consequences when seeking to use Sequoia reflectances for the retrieval of biochemical parameters.
It should also be noted that the ELM does not perform better than the single panel calibration for the grass and concrete reference targets within the survey flight. As the ELM was derived from panels imaged at a lower altitude and slightly different sun-sensor geometry, the difference is likely BRDF related.

![Graphs showing reflectance values for different calibration methods and panels](image)

**Figure 3:** Reflectance values as derived from multispectral imagery using different calibration methods for five reference panels of different reflectance as well as a pseudo-invariant concrete surface (blue) and grass meadow (green). Results are plotted by band: a) Green, b) Red, c) Red edge and d) NIR bands.

<table>
<thead>
<tr>
<th>Calibration Method</th>
<th>RMSE GREEN (SD)</th>
<th>RMSE RED (SD)</th>
<th>RMSE RED-EDGE (SD)</th>
<th>RMSE NIR (SD)</th>
<th>RMSE all bands (SD)</th>
</tr>
</thead>
<tbody>
<tr>
<td>GreyCard</td>
<td>0.0113 (0.0014)</td>
<td>0.0067 (0.0017)</td>
<td>0.0159 (0.0022)</td>
<td>0.0274 (0.0031)</td>
<td>0.0154</td>
</tr>
<tr>
<td>Spectralon (SRT-40-050)</td>
<td>0.0213 (0.0016)</td>
<td>0.0153 (0.0019)</td>
<td>0.0206 (0.0023)</td>
<td>0.0288 (0.0031)</td>
<td>0.0215</td>
</tr>
<tr>
<td>ELM</td>
<td>0.0036 (0.0016)</td>
<td>0.0022 (0.0019)</td>
<td>0.0056 (0.0024)</td>
<td>0.0032 (0.0034)</td>
<td>0.0037</td>
</tr>
</tbody>
</table>

*Table 2: RMSE per band and calibration method between the measured reflectance and the reference reflectance for the reference panels. In brackets: Standard deviations over the pixels of each panel.*

A visual comparison of the green band surface reflectance (HCRF) results for a subset of the actual surveyed scene is displayed in Figure 4. It is evident that there are slight offsets depending on the calibration method used, with the highest
values for the Spectralon (SRT-40-050) calibrated product and the ELM resulting in generally lower values, even slightly negative for the dark cast-shadow region.

![Figure 4: HCRF values in the green wavelength band for a subset of the scene showing an oak tree canopy and cast shadow for a) the standard ELM, b) simplified ELM with grey card and c) simplified ELM with Spectralon panel.](image)

3.2 Experiment 2

The derived pseudo-irradiance values for each flight are visualised in Figure 5. The recorded camera positions for F2 appear more unequally spaced as opposed to F1. As the wind speeds and directions were similar for both flights, this is likely a camera related issue appearing due to continuous operation. As frontal image overlap was very high, this issue is not expected to significantly impact the results.

The surface reflectance for F1 and F2 and the two different calibration approaches are displayed for the NIR wavelength band in the area of interest in Figure 6. As the single panel calibration image was acquired during overcast conditions, HCRF is greater than 1 for areas imaged during direct irradiance conditions if irradiance is not considered. If measured irradiance is considered, illumination-based anomalies are reduced in magnitude. Visual inspection of F1 shows that there are no large differences apparent due to more stable irradiance conditions. Comparing the F2 reflectance maps to F1, it is clear that considerable variations due to irradiance are still present, even when irradiance information is included in the calibration. The spatial distribution of brighter and darker areas is however different.
Figure 5: Spatial representations of the pseudo-irradiance values captured per image in the green band for the two drone acquisitions. Left: Flight 1 (F1), Right: Flight 2 (F2). The black dots represent the GPS position recorded for each captured image.

Figure 6: Surface reflectance in the NIR band for the consecutive acquisitions F1 and F2, with and without accounting for measured irradiance information.
The values of indices derived using band-ratios are impacted due to wavelength-dependent differences in apparent reflectance depending on direct or diffuse illumination conditions. As is evident in Figure 7 and Table 3, the mean index values differ considerably between F1 and F2 due to illumination dependent differences. The difference between mean values is greatly reduced by compensating for irradiance and the value spreads are reduced, also clearly visualised in Figure 8. However, a paired t-test showed that the differences between means was still significant for both NDVI and CHL values (p<0.05). The standard deviation of CHL values within the more variable acquisition F2 are still almost double those of F1. This means it is likely that index variations even within irradiance-compensated products may mask actual physiologically based differences between palm canopy reflectances.

Table 3: Means and standard deviations of per-palm index values for the two acquisitions F1 and F2, with and without irradiance compensation

<table>
<thead>
<tr>
<th></th>
<th>F1</th>
<th>F2</th>
<th>F1 corr.</th>
<th>F2 corr.</th>
</tr>
</thead>
<tbody>
<tr>
<td>NDVI mean</td>
<td>0.892</td>
<td>0.846</td>
<td>0.903</td>
<td>0.898</td>
</tr>
<tr>
<td>NDVI SD</td>
<td>0.013</td>
<td>0.025</td>
<td>0.011</td>
<td>0.017</td>
</tr>
<tr>
<td>CHL mean</td>
<td>0.870</td>
<td>0.308</td>
<td>0.986</td>
<td>0.928</td>
</tr>
<tr>
<td>CHL SD</td>
<td>0.164</td>
<td>0.214</td>
<td>0.125</td>
<td>0.241</td>
</tr>
</tbody>
</table>

Figure 7: Boxplots of palm median NDVI (left) and CHL (right) values for acquisitions F1 and F2, with and without irradiance compensation (corr.). The bold central line represents the median, the lower and upper boundaries of the boxes are the 25th and 75th percentile respectively and the whiskers extend to the furthest data points within 1.5 times the interquartile range, outliers beyond are omitted.
4 DISCUSSION AND CONCLUSIONS

In this study, we firstly assessed the accuracy of drone-based MCA data derived surface HCRF for two different calibration approaches and reference panels (experiment 1) as well as the impact of illumination variations on HCRF derived vegetation indices and the performance of corrections using irradiance data (experiment 2).

4.1 Experiment 1

We found that the standard ELM based method can provide surface reflectance factor estimates which are more consistent with in-situ measured values, if acquired at the same altitude and viewing geometry. The ELM can compensate for the influence of aerosols and water vapour between the ground and sensor given a uniform distribution of these throughout the imaged scene, whereas a calibration based on a single reference panel imaged on the ground would be susceptible to this bias. The ELM also has the advantage that the panels are imaged with the same viewing geometry as ~50% of the scene (flight lines of same orientation). However, as in this study the panels were imaged at a significantly lower altitude than the rest of the scene (45 m difference), atmospheric effects and viewing geometry will be slightly different and indeed for the two targets measured within the survey flight, the performance of the same standard ELM and single-panel calibration was comparable.

To enable the ELM at flight altitude, significantly larger panels would be required. For practical reasons, this is commonly solved by using tarpaulins \(^{18}\), though these cannot be levelled and may introduce variations in reflectance due to BRDF effects.

\[ \text{Figure 8: Per palm CHL and NDVI values for two replicate flights of varying illumination. Red: Irradiance considered during reflectance calibration, Blue: Irradiance not considered during reflectance calibration.} \]
When using the standard ELM a divergence in areas of low reflectance (cast shadows) was observed when the ELM intercept is not constrained to zero. Assuming adjacency effects are indeed negligible for the panels imaged at lower altitude and dark current offset is constant, the cause of this phenomenon is unclear but may be due to non-linearity in sensor response for very low incoming radiance.

There did not appear to be a large difference when considering the use of the 42.5% Spectralon panel for single-panel calibration as compared to a much more affordable Kodak grey card. Although the Spectralon panel can be expected to exhibit more Lambertian reflectance properties than the grey card, it should be noted that carbon-dosed Spectralon has been found to be slightly less Lambertian than purely white panels \(^{19}\). It should however be stressed that our study did not investigate degradation of actual reflectance versus the calibrated reflectance values over time which has been shown to be larger for lower quality panels such as the Kodak grey card used here \(^4\).

### 4.2 Experiment 2

Experiment two showed that corrections for irradiance variations based solely on irradiance sensor information can improve the consistency of results acquired during overcast conditions with varying diffuse illumination but cannot fully compensate for the impact of direct/diffuse irradiance variations. This is due to the point-based irradiance measurement representative for the image not being representative for the actual top-of-canopy irradiance of the imaged scene due to a large spatial extent being imaged, often not perfectly nadir due to non-gimbaled acquisition. Cloud shadows and gaps can therefore lead to significant variations in derived indices which is superimposed on natural variation within the scene. This is true even for the NDVI which is considered more robust to wavelength dependent variations \(^{17}\).

From the two experiments presented here, we conclude that standard ELM calibrations of pseudo-radiance orthomosaics generated using black-box processing software remain the most reliable way of deriving true HCRF from MCA drone imagery, however the deployment of panels of sufficient size and Lambertian reflectance is limiting in practice. Reference panels imaged in-flight are still highly recommended either for calibration or accuracy assessment of any study seeking to make use of individual band HCRF information. Furthermore, in order to generate spatially consistent HCRF products, irradiance conditions should be as homogenous as possible and any greater variations will introduce errors which cannot be adequately compensated using drone-based irradiance measurements. There is therefore a need for scientists working with multispectral drone data to be aware of these limitations and transparently report illumination conditions during acquisitions, ideally through concurrent measurements.

### 5 ACKNOWLEDGEMENTS

D. Fawcett and K. Anderson received funding from the European Union’s Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie Grant Agreement No 721995.

We would like to acknowledge the help of Khoon Lip Kho of the Malaysian Tropical Peat Research Institute for facilitating the work performed in Sarawak, Malaysia and Sarawak Oil Palms Berhad for hosting and making the site available.
6 REFERENCES


