Cost-effective non-metric photogrammetry from consumer-grade sUAS: implications for direct georeferencing of structure from motion photogrammetry

Patrice E. Carbonneau1* and James T. Dietrich2
1 Department of Geography, Durham University, Durham, UK
2 William H. Neukom Institute for Computational Science, Dartmouth College, Hanover, NH USA

Received 18 December 2015; Revised 27 July 2016; Accepted 28 July 2016

*Correspondence to: Patrice E. Carbonneau, Department of Geography, Durham University, Lower Mountjoy Site, South Road, Durham, UK. E-mail: patrice.carbonneau@durham.ac.uk

ABSTRACT: The declining costs of small Unmanned Aerial Systems (sUAS), in combination with Structure-from-Motion (SfM) photogrammetry have triggered renewed interest in image-based topography reconstruction. However, the potential uptake of sUAS-based topography is limited by the need for ground control acquired with expensive survey equipment. Direct georeferencing (DG) is a workflow that obviates ground control and uses only the camera positions to georeference the SfM results. However, the absence of ground control poses significant challenges in terms of the data quality of the final geospatial outputs. Notably, it is generally accepted that ground control is required to georeference, refine the camera calibration parameters, and remove any artefacts of optical distortion from the topographic model. Here, we present an examination of DG carried out with low-cost consumer-grade sUAS. We begin with a study of surface deformations resulting from systematic perturbations of the radial lens distortion parameters. We then test a number of flight patterns and develop a novel error quantification method to assess the outcomes. Our perturbation analysis shows that there exists families of predictable equifinal solutions of $K_1 - K_2$ which minimize doming in the output model. The equifinal solutions can be expressed as $K_2 = f(K_1)$ and they have been observed for both the DJI Inspire 1 and Phantom 3 sUAS platforms. This equifinality relationship can be used as an external reliability check of the self-calibration and allow a DG workflow to produce topography exempt of non-affine deformations and with random errors of 0.1% of the flying height, linear offsets below 10 m and off-vertical tilts below 1°. Whilst not yet of survey-grade quality, these results demonstrate that low-cost sUAS are capable of producing reliable topography products without recourse to expensive survey equipment and we argue that direct georeferencing and low-cost sUAS could transform survey practices in both academic and commercial disciplines. Copyright © 2016 John Wiley & Sons, Ltd.

KEYWORDS: sUAS; UAV; Structure-from-Motion; SfM; photogrammetry; direct georeferencing; camera calibration; point clouds

Introduction

The staggering uptake of small Unmanned Aerial Systems (sUAS), commonly referred to as drones, has led some to coin the term ‘the Drone Age’ in reference to our era (Economist, 2015). In parallel, the integration of Structure-from-Motion (SfM) and multi-view stereo (MVS) algorithms into the standard workflow of digital photogrammetry has led to a series of software products that can restitute topography from imagery with an unprecedented level of automation and ease (James and Robson, 2012; Westoby et al., 2012; Forstal et al., 2013). Several authors have already demonstrated that imagery acquired from sUAS platforms can deliver high quality topographic survey data (Niethammer et al., 2012; Hugenholtz et al., 2013; Bemis et al., 2014; Immerzeel et al., 2014; Clapuyt et al., 2015; Elnser et al., 2015; Ryan et al., 2015; Turner et al., 2015; Woodget et al., 2015). The appearance of consumer-grade sUAS with imaging sensors that approach the quality of professional and scientific digital cameras is therefore precipitating a fundamental shift in topographic mapping whereby individuals or small organizations in need of such data no longer depend on national mapping agencies or geospatial/survey companies. Whilst some progress has been made towards lowering the cost of terrain mapping via digital photogrammetry (e.g. Carbonneau et al., 2003), the uptake has largely remained contained to academic circles with the commercial topographic sector currently dominated by laser-ranging technology. However, with low-cost drones and affordable SfM-photogrammetry software, photogrammetry is gaining and may return as the dominant method of topography production and become as common a tool as standard photography.
Currently, the combination of a consumer-grade sUAS equipped with a high-definition camera and a free SfM-photogrammetry software, such as 123D Catch by Autodesk or Microsoft Photosynth, can produce virtual three-dimensional (3D) objects that are suitable for visualization or 3D printing. However, the production of topographic data and the associated orthoimagery for the purpose of formal mapping or any other scientific application is significantly more challenging and has much more stringent data quality requirements. The 3D data initially produced by SfM are not scaled and oriented to any real dimensions or directions (James and Robson, 2012; Westoby et al., 2012; Forstnad et al., 2013; Javemic et al., 2014). Therefore, SfM photogrammetry requires generally expensive survey equipment [e.g. real-time kinematic global positioning system (RTK-GPS) or total station] to provide accurate (i.e. small mean error) and precise (i.e. small standard deviation of error) real-world map-coordinates of objects in the scene [e.g. ground control points (GCPs) or targets]. These points are used to scale, rotate, and translate the initial model to map coordinates. However, errors present in the survey data propagate through the transformation and result in errors of scaling, rotation, and translation as illustrated in Figure 1. In addition to these linear errors, detailed investigations have revealed that systematic doming deformations can often be present in the final topographic outputs for both standard photogrammetry (Wackrow and deformed) and SfM-MVS (Conrady et al., 2013; Javemic et al., 2014; Woodget et al., 2015; Dietrich, 2016). These deformations are caused by optical lens distortion (Figure 2). Figure 2a illustrates how lens distortion can warp images. Under ideal circumstances, a lens would produce an image with a regular orthogonal grid pattern preserves straight lines and right angles. In practice, lens design and aberration effects warp the conformal image projection and result in non-right angles and curved lines with either barrel or pin-cushion distortion patterns. Furthermore, many compact lens systems that are increasingly prevalent in small-format and mobile cameras display complex patterns where the distortion is not a monotonic function of radial distance from the centre of the image. When the lens distortions from multiple images are combined and allowed to propagate through the SfM-MVS process the effect on the final topographic outputs is generally seen as either a doming or dishing effect (Figure 2b). Correcting the lens-distortion effects has been the focus of a very large body of work in both photogrammetry and computer vision (Brown, 1966, 1971; Fraser, 1997; Clarke and Fryer, 1998; Luhmann et al., 2014). Use of the Brown–Conrady lens distortion model (Table I) is now standard practice in all forms of photogrammetry. Furthermore, the parametrization of this model is now performed by so-called ‘self-, automatic-, or on-the-job self-calibration’ algorithms which automatically parametrize the Brown–Conrady model (i.e. calibrating the camera/lens combination) from the image dataset intended for 3D reconstruction (Fraser, 1997; Luhmann et al., 2014).

High-quality SfM-photogrammetry, therefore, requires both a correct scaling, rotation, and translation of a raw 3D point cloud and an accurate parametrization of the Brown–Conrady lens distortion model. The dominant approach is the acquisition of centimetric quality ground control data (GCPs) with professional survey equipment (e.g. RTK-GPS or total station). The photogrammetric software (standard or SfM) can then use ground control data for both georegistration and camera calibration. However, the requirement for GCPs to be acquired with survey-grade equipment remains the largest barrier, both in terms of time and cost, to the wider uptake of sUAS-based photogrammetry. The alternative is to develop so-called ‘Direct Georeferencing (DG)’, whereby the photogrammetric solution is determined with precise and accurate knowledge of the camera positions (X, Y, and Z) and orientation (pitch, roll, and yaw) at the time of image acquisition (Turner et al., 2014). This methodology is routinely used in the case of full-sized airborne photography and LiDAR (light detection and ranging) surveys (without the camera calibration requirement) and some authors have started reporting on its application to sUAS (Turner et al., 2014).

The advantages of DG are clear: by removing the need for ground control, professional survey equipment is no longer required which could dramatically reduce the cost of topographic surveys. Furthermore, with DG, ground access to the survey area would not be required, which would facilitate high-quality topographic monitoring in hazardous or inaccessible areas. We therefore argue that the development and wider uptake of cost-effective photogrammetry from sUAS platforms that are capable of producing outputs suitable for a range of mapping and technical applications is reliant on the further development of DG. To a certain extent, this has already begun. In recent years, there has been an explosion of sUAS options made available to consumers and researchers, both fixed-wing and rotorcraft. They range from higher priced ready-to-fly (RTF) models to lower priced do-it-yourself or ‘scratch-built’ options. As discussed further in this paper, these drones have many desirable features for scientific mapping and they have enabled an emergence in consumer-grade photogrammetry, colloquially called ‘3D mapping’. However, much of this activity is not informed by photogrammetry and the specific challenges of data quality and camera calibration quality in the absences of GCPs have not been addressed in the consumer market and only researched by a few authors in academic circles (Turner et al., 2014). DG of SfM has been demonstrated from manned aircraft (Nolan et al., 2015). However, the reliability of the DG approach from sUAS and its suitability as a fully-fledged topographic survey method is yet to be demonstrated.

In this paper, we deploy two popular consumer-grade sUAS models and we examine their ability to produce topography with a DG workflow. Our approach is informed by core photogrammetry concepts and particular attention was paid to camera calibration issues: Perturbation analysis of lens distortion and corresponding doming deformations were conducted within a space of the $K_1$ and $K_2$ lens distortion parameters (Table I). Furthermore, we develop a new approach to error assessment which is tailored to the SfM-photogrammetry workflow and explicitly determines errors in translation, rotation, scale, and surface noise and demonstrates that whilst some form of limited ground truthing is still required,
survey-grade ground control is not required to achieve a satisfactory data quality.

**Methods**

sUAS platforms

We deployed a Phantom 3 Professional (P3P) and an Inspire 1 (I1) both manufactured by DJI Inc. (Shenzhen, China). These popular quadcopters have many features of interest for scientific applications. The P3P and the I1 both have an integrated camera (Model FC300 on the P3P and Model FC350 on the I1) mounted on a three-axis gimbal that stabilizes the camera, absorbs vibrations, and compensates for the rotational motion of the quadcopter. The integrated cameras both use a wide-angle rectilinear lens and thus avoid the heavy distortions common with the fish-eye lenses employed in several drone and camera models (e.g., GoPro cameras or the DJI Phantom 2-series). Both the P3P and I1 cameras can acquire 12-megapixel still imagery in raw (DNG) or JPEG formats. For navigation and flight stabilization, the internal consumer-grade GPS uses both the GNSS and GLONASS systems, increasing the number of satellites used in the GPS position determination. DJI’s specifications report a GPS accuracy of ≈2.5 m in X–Y. These positions are automatically exported to the EXIF metadata for each image in WGS84 latitude and longitude thus providing a location stamp (geotag) for each image. Finally, the low-cost of the P3P (£1200) and the I1 (£2400) make them accessible and limits the impact of a total loss in the event of a crash, allowing for flights in risky environments such as volcanos or large flood-stricken areas.

**Table I.** Camera calibration parameters in the Brown–Conrady model.

<table>
<thead>
<tr>
<th>Parameters</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>$F_x$, $F_y$</td>
<td>Focal length of the camera in X and Y dimensions.</td>
</tr>
<tr>
<td>$C_x$, $C_y$</td>
<td>Principle point of the image (x and y image coordinates)</td>
</tr>
<tr>
<td>$K_1$</td>
<td>Second-order radial distortion</td>
</tr>
<tr>
<td>$K_2$</td>
<td>Fourth-order radial distortion</td>
</tr>
<tr>
<td>$K_3$</td>
<td>Sixth-order radial distortion</td>
</tr>
<tr>
<td>$P_1$, $P_2$</td>
<td>$X(P_1)$ and $Y(P_2)$ tangential distortion</td>
</tr>
</tbody>
</table>

Brown–Conrady distortional model (Brown, 1966; Heikkila and Silven, 1997):

$$
\begin{align*}
\hat{x} &= (1 + K_1 r^2 + K_2 r^4 + K_3 r^6) x + 2P_1 xy + P_2 (r^2 + 2x^2) \\
\hat{y} &= (1 + K_1 r^2 + K_2 r^4 + K_3 r^6) y + 2P_1 xy + P_2 (r^2 + 2y^2) + 2P_2 xy
\end{align*}
$$

$$
\tau = \sqrt{(x - C_x)^2 + (y - C_y)^2}
$$

where: $x$, $y$ are the distorted pixel coordinates; $\hat{x}$, $\hat{y}$ are the undistorted pixel coordinates.

**Camera calibration experiment**

The optical components of any real lens system do not transmit light rays from the object scene to the imaging sensor in a perfectly linear manner (Wolf et al., 2014). In order for SfM, or any photogrammetric method, to recover accurate 3D data defining and correcting lens distortions are critical (Zhengyou, 1996; Heikkila and Silven, 1997). Lens distortions are typically generalized with two components: symmetric radial distortions and decentring (tangential) distortions (Förstner et al., 2013). The Brown–Conrady lens model was developed in order to correct these lens distortions. In Table I, we can see the form of the model which is based on Taylor expansions of radial, $K_n$, and tangential $P_n$, distortion terms. The radial distortion parameters ($K_1$, $K_2$, and $K_3$) are associated with barrel and pincushion distortion patterns (Wolf et al., 2014) (Figure 2) and are combined to create a two-dimensional (2D) polynomial representation of the lens distortions. The $P_n$ terms quantify the decentring distortions, offsets of the radial distortion parameters.
from the centre of the image. Fraser (2013) states that there is a temptation to use the full 10-parameter form of the model to achieve the highest possible accuracy. However, the two most critical terms in the model are the focal length (f) and K1, followed by K2 and the principal point coordinates (Cx, Cy) (Förstner et al., 2013; Wolf et al., 2014). For ‘medium accuracy’ applications, K1 is sufficient, but for higher accuracy applications, especially those utilizing wide-angle lenses (e.g., those used in modern, compact camera systems), K2 and possibly K3 are required (Förstner et al., 2013). For K1, P1, and P3 to be solved accurately the image network needs to be highly redundant with a very strong, convergent geometry (Wolf et al., 2014), which is often only possible in highly controlled environments. It is recognized that the K1 parameters can be highly correlated, due to the nature of the polynomial model (Ziemen, 1986; Fraser, 1997). Fraser (1997) suggests that statistical tests can be conducted on the parameters to determine if additional K1 parameters result in a statistically significant difference in the radial distortion profile. However, the K2 parameters are not strongly coupled to the other parameters in the model or with the external orientation parameters and therefore an overparameterization (additional K1 terms) will still yield a valid distortion profile (Fraser, 1997). Additionally, P1 and P2 are highly correlated with the principal point coordinates and the associated displacements are typically small compared to the K-terms. If P1 and P2 are supressed (i.e. set to zero) the errors can be absorbed by Cx, Cy (Fraser, 2013; Förstner et al., 2013). Förstner et al. (2013) also states that the distortions associated with P-terms are ‘universally ignored in analytical photogrammetry’.

When errors associated with radial image distortion are allowed to propagate through the photogrammetric process, it has been established that the final topographic model can contain non-linear deformations that take the form as doming or dishing of the land surface (Figure 2b) (Wackrow and Chandler, 2008; James and Robson, 2014). We therefore begin with an experiment aimed at determining the optimal radial distortion parameters for the P3P and I1 lenses. We used Photoscan Professional (version 1.1) for all the processing described in this work. However, this software expresses the earlier mentioned distortion parameters in focal length units according to:

\[ K_1(\text{focal units}) = K_1(\text{pixel units})^2 \]  
\[ K_2(\text{focal units}) = K_2(\text{pixel units})^3 \]  
\[ K_3(\text{focal units}) = K_3(\text{pixel units})^5 \]

Where f is the focal length in pixel units. The expression of \( K_0 \) in focal length units yield easily manageable numbers which can be expressed without scientific notation, but it hinders cross-comparability of calibration results with different focal lengths. Therefore, from this point onwards, we will systematically report \( K_0 \) values in pixel units as converted from Equations (1)-(3).

To identify the correct distortion parameters for our lenses, we used a systematic perturbation approach with the objective of iteratively finding the optimal \( K_1 \) and \( K_2 \) values that minimized the distortions in a flat wall test site. The sUAS were operated hand-held without propellers. By walking in front of a flat wall on the campus of Durham University (for the P3P) and Dartmouth College (for the I1) at a distance of roughly 10 m, a series of 24 images were acquired as a series of 12 convergent viewing pairs. For each pair, the optical axis of the camera intersects the wall at \( \pm 45^\circ \). Our only assumption is that the wall is flat and, in keeping with the objective of DG, no other ground validation is used. Based on our other research and recreational photogrammetry experiments and data collections, the range of \( K_1 \) and \( K_2 \) values was set as \( K_1 \in [-3.7, -1.9] \times 10^{-8} \) pixels and \( K_2 \in [1.7, 7] \times 10^{-15} \) pixels converted as per Equations (1)-(3) with a focal length of 2320.06 pixels for the P3P and 2326.07 for the I1.

Within Photoscan, the 24 flat wall images were processed in order to establish initial camera calibration values. Using the Python scripting API (application program interface) for Photoscan we were able to produce and process a total of 6232 photogrammetric blocks with the same imagery but with fixed calibration parameters spanning the full \( K_1 - K_2 \) space while keeping the calibration values for the focal length, principal point offset and \( K_3 \) as per the initial self-calibration. Tangential distortion and \( K_4 \) were not used in this experiment. The resulting point clouds were exported automatically as part of the Python processing script. In MATLAB, each model was centred and normalized to the same dimensions and a 3D second-order polynomial fit in \((x,y)\) was calculated for each point-cloud data. Given that the test wall is flat, any significant doming distortions in the point clouds are detected in the second-order terms of the polynomial regression. In this case, the wall was longer in the x direction and therefore we expect significant doming to be detected in the x^2 term of the polynomial. Optimal calibration parameters will be those where these second-order terms are minimized (\( \equiv 0 \)). The compiled results of the experiments were plotted as a matrix to display the data as a function of \( K_1 \) and \( K_2 \).

Lens distortion modelling

Our personal experience with camera calibration has shown that for a given camera, the outputs of self-calibration are never identical even in cases where imagery is acquired on the same day, under similar conditions and where external validation indicates that the final topographic model is of high quality. In order to further examine similarity in the outputs of the camera calibration as observed in the perturbation experiments, we conducted additional simulations of lens perturbations. We used a Monte Carlo framework (i.e. randomly generated perturbations) implemented in MATLAB. This random Monte Carlo approach is less computationally intensive than a full examination of the parameter space. To start, we take the optimal \( K_1 \), \( K_2 \) and \( K_3 \) calibration parameters as determined by self-calibration in the flat wall experiment described earlier and we calculate the associated displacement with the Brown–Conrady model (Table I), in pixels, over a profile of 2000 pixels (the positive horizontal axis of our photographs). Then a Monte Carlo approach is used to randomly generate one million combinations of \( K_1 \) and \( K_2 \) once again in the interval \( K_1 \in [-3.7, -1.9] \times 10^{-8} \) pixels and \( K_2 \in [1.7, 7] \times 10^{-15} \) pixels (as earlier). For each of the parameter combinations the pixel displacement profile is re-calculated. We then take the maximum difference, irrespective of location in the distortion profile, between the simulated \( K_1 - K_2 \) combination and the optimal \( K_1 - K_2 - K_3 \) calibration. Finally, the maximal difference results from the one million Monte Carlo samples are interpolated to a regular grid for analysis and display.

Airborne surveys and validation

Additional experiments were conducted with the P3P at two sites in county Durham, UK. First, a fallow field in the village of Lancaster was used as the main site (Site A). This triangular field is ringed by trees and it is characterized by a curved
topography which is sloping in an eastwards direction. The
field is fallow with wild grasses and flowers grown to a height
of 10 to 40 cm. Second, a sports field on the grounds of
Durham University was used (Site B). This area was flat and
had trimmed lawn. The northern limit of the area is set by a
flood levee approximately 1 m high. Figure 3 shows an aerial
view of each site. The topography of both sites was established
with an RTK-GPS survey using a Leica 1200 model RTK rover
and base pair. For each site, RTK-GPS points were acquired
with the objective of establishing the overall shape of the area.
Instead of working on a gridded basis, care was taken to
capture the outer bounds of the study sites and break-lines in
the local topography. The GPS base-stations were set to relative
positions but post-processed with added data from a permanent
base station located at Newcastle Airport. The results of the
survey were two sparse GPS point datasets that captured the
shape of each study site.

Flight patterns and image acquisition

Based on the work of James and Robson (2014) and Fonstad
et al. (2013) there is a consensus view that image acquisition
geometry has an impact on the quality of the outputs with
convergent views and multiple flight altitudes to be preferred.
We therefore designed an intuitive set of flight patterns based
on reasonable combinations of altitude, camera orientation,
and flight direction that are expected to illustrate both good
and bad performance of the SfM-photogrammetry process and
test the potential of DG. The flights were grouped in two
experiments both aiming to test the effect of flight pattern and
image acquisition on final model geometry (detailed in
Table II). We experimented with a combination of NADIR and
convergent imagery (James and Robson, 2014). Four flight
patterns were tested: (1) NADIR imagery from an altitude of
60 m; (2) NADIR imagery from an altitude of 60 and 80 m; (3)
NADIR imagery combined with eight convergent-view images
at an altitude of 60 m; (4) NADIR imagery at 60 and 80 m
combined with eight convergent view images acquired at
60 m. It was found that the P3P had sufficient battery life to
acquire all the imagery for these four flight patterns in one
flight. Flights began with the camera at NADIR at two altitudes
and ended with the series of oblique view images. In the
analysis stage, the images were grouped according to the four
flight patterns for both sites thus yielding eight ‘flights’ labelled
A1–A4 and B1–B4. Finally, at the time of our experiments, the
DJI firmware for both the P3P and 11 recorded the relative
elevation from the launch point to the EXIF data of each
photograph. Before every flight, we needed to establish a zero
elevation datum for the flight by taking a single image on the
ground from the launch position in order to get the GPS
position corresponding to the launch point.

Pre-processing

For all eight flights, the raw RGB images (DNG format) were
converted to 48-bit TIFF imagery (16-bit per band) with Adobe
Photoshop elements. No image equalization or other such
adjustments were applied. We imported the initial TIFF images
into MATLAB, removed the geotag (latitude/longitude) from the
EXIF data, converted and output the coordinates into UTM
(Zone 30 North, WGS84) to a text file, which could be
imported into Photoscan Pro. In addition, the X–Y coordinates
of the launch point were used to get the absolute elevation from
a hydrologically corrected SRTM (Shuttle Radar Topography
Mission) dataset available from the United States Geological
Survey (USGS). This value was added to the relative elevation
data recorded in the image metadata thus resulting in each
image defined by a 3D position in UTM coordinates with
elevations above the WGS 1984 ellipsoid.

TIFF images are then imported into Photoscan Pro in separate
‘chunks’ which act as distinct processing blocks. The converted
camera station coordinates were imported providing an
immediate geographic reference for each block. Based on the
DJI manual, which states that the on-board GPS is accurate to
2.5 m in X–Y, the camera position accuracy setting, in
Photoscan, was set to a conservative value of 5 m. Each image
block was reconstructed using the high quality image-
alignment setting followed by a dense reconstruction at a
low-density setting.

Figure 3. Study Sites A and B near Durham UK, with profiles locations used in Figures 7 and 8. The colour difference in the Site A image (around α)
was due to blending errors in orthophotograph for this shadowed area.
Error and quality assessment for 3D point clouds and topography data typically relies on root mean square error (RMSE) values when compared to some ground truth-value (ASPRS, 2015; Whitehead and Hugenholtz, 2015). We argue that these generic metrics do not provide a full description of the possible errors associated with SfM-photogrammetry. A single root mean square (RMS) statistic cannot explicitly identify systematic patterns such as tilt or non-affine warp. Mapping the spatial distribution of error at checkpoints can be useful in identifying these patterns, but often low checkpoint density does not capture the full extent of any systematic distortions. In the absence of a dense network of survey-grade GCPs, the georeferencing process used in SfM-photogrammetry relies on a rigid seven-parameter transformation, errors in the ground control or camera position can propagate as errors of position, orientation, and/or scale in the model (Figure 1). Each of these errors has impacts with respect to specific applications. Position/translation errors can affect change detection studies by creating false horizontal or vertical offsets. Orientation errors, especially off-vertical rotations, will affect gravity-dependent models (i.e. flow models) by producing incorrect slopes and flow directions. Scale errors will also influence change detection and volume calculations, again by creating false horizontal or vertical offsets. A single RMS statistic, even if accompanied by a full error distribution, is therefore not a powerful diagnostic tool when deciding if a topographic dataset is suitable for any given application. A more process-specific error assessment approach is therefore required. We propose the following error model:

\[ P_{\text{SfM}} = M_{\text{true}}P_{\text{true}} + \eta \]  

(4)

Each of the three terms in Equation (4) is a point cloud: \( P_{\text{true}} \) is a reference point-cloud giving the true (i.e. ground truth) representation of the surface. The term \( P_{\text{SfM}} \) is the point cloud calculated by SfM-photogrammetry. The matrix factor \( M_{\text{true}} \) denotes the affine rigid-body seven-parameter transform needed to scale (one parameter), rotate (three parameters: \( R_x, R_y, R_z \)), and translate (three parameters: \( T_x, T_y, T_z \)) \( P_{\text{true}} \) in order to match \( P_{\text{SfM}} \). The value \( \eta \) is a non-linear, non-rigid, error term which can be seen as a quasi-random field produced by noise in the SfM process. We can therefore use the standard deviation of \( \eta \) as a measure of precision (scatter) of \( P_{\text{SfM}} \). This will be reported as a single number and noted as \( \eta_{\text{SfM}} \).

To assess errors in the SfM reconstructions we used differential global positioning system (dGPS) survey data and the open-source point cloud processing software Cloud Compare (Girardeau-Montaut, 2016). The first step in the process is establishing the parameters for \( M_{\text{true}} \) for all eight experiments. We duplicated each model (A1–A4 and B1–B4) in Photoscan and using the point cloud editing functions, edited each model to clear all points except those around the periphery of the sites where dGPS point density was the greatest. For Site A, only those points located along the unvegetated and narrow footpath that circumscribes the field were kept. For Site B, only the narrow concrete border (3 cm in width) of the paved footpath and the bottom and top of the flood defence ridge were preserved. These 'hollow' clouds were imported into Cloud Compare along with the matching dGPS data. Then, automated cloud-to-cloud registration, an iterative closest point algorithm, was employed to calculate a transformation matrix that could transform the dGPS data \( P_{\text{true}} \) to match the hollow \( P_{\text{SfM}} \) clouds. The resulting transformation matrix is \( M_{\text{true}} \) in Equation (1). The accuracy and precision of the automated cloud-to-cloud registration procedure were evaluated separately for Sites A and B and we use the residuals of points after the adjustment as an indication of fit quality. After automated registration for Site A, we obtain a mean residual magnitude of 5.2 cm with a standard deviation of 23.4 cm. For Site B, we obtain a mean residual magnitude of 0.1 cm and a standard deviation of 3.1 cm. The higher errors for Site A are to be expected due to the presence of tall grasses which will add error in the comparison of a bare earth digital surface model (DSM) derived from dGPS to a digital terrain model (DTM) derived from photogrammetry.

Next we calculated \( \eta \) with the full point clouds. In practice, we calculate \( \eta \) as the differences between \( P_{\text{true}} \) and \( P_{\text{SfM}} \) after the application of \( M_{\text{true}} \). The transformation matrix derived from the first step was applied to the full dGPS dataset and we constructed a Delaunay mesh of the dGPS points. Then the full SfM point cloud was imported into Cloud Compare and a cloud-to-mesh distance was calculated. The advantage of using the cloud-to-mesh function is that distances between the transformed \( P_{\text{true}} \) and \( P_{\text{SfM}} \) are calculated along the Z-axis and the sign of the difference is preserved. Positive differences are above \( P_{\text{true}} \) and points with negative differences are below \( P_{\text{true}} \). The topography at our sites was relatively flat allowing us to use the simpler cloud-to-mesh distances along the Z-axis. In more complicated terrain with vertical surfaces and/or overhangs, other differencing algorithms, such as Multiscale Model to Model Cloud Comparison (M3C2), could be used to generate the difference map (Lague et al., 2013). The differences are stored as an additional field in the point cloud, and we calculate precision by taking the standard deviation of \( \eta \). Furthermore, we extracted profiles from the difference clouds (see Figure 3 for profile locations) and used in conjunction with the spatial distribution of cloud-to-mesh differences we were able to check for the presence of the doming deformation associated with errors in the camera calibration process. Our main objective here was to test if a hypothesized degradation of camera calibration parameters would result in the now famous doming deformation (Wackrow and Chandler, 2008, 2011; James and Robson, 2014; Woodget et al., 2015). In this case, the reference surface could not be taken as the dGPS surface. The presence of tall grasses on Site A hides any dome since the dGPS data represents the ground and the photogrammetric point clouds represent the top of the vegetation, which is a variable 10 cm to 40 cm above the bare earth. Therefore, the profiles were based on what was assumed as the best possible reference for the top of the vegetation: flights A4 and B4. In all cases, we corrected for translation, rotation, and scaling errors with the application of the appropriate \( M_{\text{true}} \) transformation matrix (as described earlier) prior to differencing and profile extraction.

<table>
<thead>
<tr>
<th>Site A</th>
<th>Site B</th>
<th>Flight pattern</th>
</tr>
</thead>
<tbody>
<tr>
<td>A1</td>
<td>B1</td>
<td>NADIR imagery from an altitude of 60 m</td>
</tr>
<tr>
<td>A2</td>
<td>B2</td>
<td>NADIR imagery from an altitude of 60 and 80 m</td>
</tr>
<tr>
<td>A3</td>
<td>B3</td>
<td>NADIR imagery combined with eight convergent-view images at an altitude of 60 m</td>
</tr>
<tr>
<td>A4</td>
<td>B4</td>
<td>NADIR imagery at 60 and 80 m combined with eight convergent view images acquired at 60 m</td>
</tr>
</tbody>
</table>
Results

Figure 4 examines the self-calibration results for the P3P camera. For the wall experiment, the self-calibration of the image block returned a focal length of 2320.06 pixels with $k_1$ of $-2.53 \times 10^{-8}$ pixels$^{-1}$, $k_2$ of $3.96 \times 10^{-15}$ pixels$^{-4}$ and $k_3$ of $-1.15 \times 10^{-22}$ pixels$^{-6}$. In Figure 4, we plot lens distortion profiles over the 2000 pixel positive horizontal axis of the image for each $K_i$ component separately. The $K_i$ component profile has a parabolic, concave-down shape. With an opposite sign, the $K_2$ component profile has concave-up shape and, especially at the edges, a similar displacement magnitude when compared to the $K_1$ component. Finally, the $K_3$ component profile can be seen to have a small contribution.

Figure 5 presents the results of the Monte Carlo lens distortion simulations for both the P3P and II lens calibrations. These graphs display a strong degree of equifinality. Along the central diagonal, it is clear that a large range of $K_1$-$K_2$ combinations can produce lens distortion profiles which match the optimal lens calibration to within displacements of 10 to 20 pixels. Further, we see zones which match to within five and even two pixels. Crucially, these zones follow the diagonal and indicate that only certain combinations of $K_1$ and $K_2$ yield equifinal solutions. Given the opposing concavity of the $K_1$ and $K_2$ components seen in Figure 4, this is a sensible result. Linear regression indicates that solutions similar to within two pixels are predicted by:

$$\text{P3P }: K_2 = -2.7364 \times 10^{-7} \times K_1 + -3.4086 \times 10^{-15}\quad (5)$$

$$\text{II } : K_2 = -2.7335 \times 10^{-7} \times K_1 + -3.5858 \times 10^{-15}\quad (6)$$

Where the domain of validity for Equation (5) is $K_1 \in [-2.56, -2.35] \times 10^{-8}$ pixels and $K_2 \in [2.97, 3.62] \times 10^{-15}$ pixels and for Equation (6) is $K_1 \in [-2.46, -2.25] \times 10^{-8}$ pixels and $K_2 \in [2.52, 3.18] \times 10^{-15}$ pixels.

Figure 6 shows the output of the perturbation experiments for the P3P and II in raster format. Similar to Figure 5, this parameter space does not show a unique minima corresponding to a unique set of $K_1$-$K_2$ values that minimize distortions. Instead, we see positive curvature (pin-cushion distortion) on the top-left and negative curvature (barrel distortion) on the bottom right. In the central diagonal, we have a hedge where a family of solutions in the $K_1$-$K_2$ parameter space that minimize doming deformations with the following fits of $K_2 = f(K_1)$, where $K_0$ is expressed in pixel units,

$$\text{P3P } : K_2 = -2.392 \times 10^3 \times K_1^3 - 18.264 \times K_1^2 - 6.924 \times 10^{-7} \times K_1 - 5.887 \times 10^{-15}\quad (7)$$

$$\text{II } : K_2 = -2.781 \times 10^3 \times K_1^3 - 19.604 \times K_1^2 - 6.900 \times 10^{-7} \times K_1 - 5.362 \times 10^{-15}\quad (8)$$

Furthermore, since each point in Figure 6 is the result of a least-squares surface fitting, we examined the fit quality of all points. For the P3P, the mean $R^2$ is 0.94, the fifth percentile is 0.75 and the median is 0.98. For the II the mean $R^2$ is 0.99, the fifth percentile is 0.99 and the median is also 0.99. Further, if we consider only the subset of solutions along the line of equifinality, we find that for the P3P the mean $R^2$ is 0.97, the fifth percentile is 0.94 and the median is 0.98. For the II, the results are the same as earlier with a mean $R^2$ of 0.99, a fifth percentile of 0.99 and the median of 0.99. The differences between the P3P results and the II results suggest an improved matching success for the II which could either be due to lighting conditions, surface texture or sensor quality. Overall these statistics indicate that the wall remained reasonably flat in the entire $K_1$-$K_2$ parameter space, especially for the subset of equifinal solutions defined by Equations (7) and (8). Figure 6 also displays points that correspond to calibration results (in pixel units) from the flight experiments with the P3P and, for the II, a sample of other data acquired at a variety of field sites. We can see that most, but not all, self-calibration outputs plot near the zero distortion line. Table III lists the self-calibration parameters returned for the eight P3P flights. We first evaluated the self-calibration by calculating the value of $K_1$ as per Equation (7) and then calculating the error, $K_{err}$, between the self-calibrated and predicted values of $K_2$. According to this metric, flights A1 and A2 were the furthest from the zero distortion line and flights A3, A4 and B2 the closest. Flights B1, B3 and B4 seem to be in an intermediate range. Equations (5) and (7) suggests that doming deformations present due to a mis-calibration of $K_1/K_2$ can be corrected with an adjustment of $K_0$ or $K_1$. Preliminary experiments with corrections derived from Equation (7) were only partial successful at removing doming. However, constraining the corrections to the zone in Figure 5 with solutions equifinal to within displacements of two pixels (i.e. using Equation (5)) was successful. When applying Equation (5), $K_1$ values self-calibrated as being outside the domain of validity were forced to the nearest value within the domain and the $K_2$ value re-calculated accordingly. Figure 7 presents both self-calibrated and adjusted deformation profiles along the $\alpha$ to $\beta$ profile in Figure 3. Deformation is calculated as a given profile differentiated from the optimally calibrated profiles A4 and B4. Figure 7 shows that self-calibration for flights A1 and B1 returned a strong doming deformation while self-calibration of A2 resulted in a dishing deformation. The adjusted calibration values eliminated both doming and dishing deformations. However, this has come at the cost of increased surface error (random noise). This level of random error, $\eta$ in Equation (4), along with other errors, are presented in Table IV. In Table IV, we note that $T_2$ (datum shifts) are large for A1 and A2, and a re-examination of Table III shows that variations in $T_2$ to be a part of a complex and non-intuitive response to focal length calibration issues. According to single image geometry, the scale of the image is directly proportional to the focal length. Therefore, if we consider the calibration results for A1 and A2, we might expect that a focal length difference of 565 out of 2539 would result in a $\approx 22\%$ difference in scale. Table IV shows that this is clearly not the case with a scaling errors of 0.2% in the case of A1. Furthermore, a re-examination of the point clouds via simultaneous display for cases A1–A2 shows that these models, with focal length calibrations ranging from 1974 to 2539 have a clear vertical offset but visible scale differences in the X-Y plane. This is illustrated in Figure 8, which shows raw elevation profiles for flights A1 and A2 taken from points along the $\alpha$ to $\gamma$ cross-section in Figure 3. The raw elevations in Figure 8 were calculated from separate digital elevation model (DEM) outputs where the exact beginning and end-points of the profiles were manually chosen.
Errors in the scaling parameter are also a potentially important source of error for surface process science applications. Finally, equipment, we have demonstrated that useable topography data can be produced with a low-cost DG approach. This promising result of 0.06, at a flying height of 60 m, would indicate an optimal precision of 0.1% (1:1000) of flying height.

**Discussion**

Our results demonstrate that DG from low-cost aerial platforms is a viable workflow for high-quality topography mapping. Predictable equifinality in the $K_1-K_2$ parameter space of calibration solutions provides a measure of external reliability capable of assessing the quality of a camera calibration without the recourse to external ground control. In optimal cases, we find that surface errors of 0.1% of flying height can be achieved. This is equivalent to James and Robson (2012) who found an almost identical performance using a control network of 45 surveyed ground control points and an SLR camera. Therefore, the novelty of this work lies in the fact that we have demonstrated and achieved a performance level previously associated to SLR cameras and survey-grade ground control using only a consumer-grade drone, with a built-in camera and an on-board GPS, delivering positions which are not processed with differential correction. By matching the data quality presented in James and Robson (2012) with consumer grade equipment, we have demonstrated that useable topography data can be produced with a low-cost DG approach.

Further consideration of the similarities with the findings of James and Robson (2012) leads to important observations. It should be noted that the code used in James and Robson (2012) did not use the high-accuracy GCPs in localized surface

Figure 5. Equifinality in $K_1-K_2$ space for the P3P and II lenses. Colour scales and contours represent the maximum deviation, in pixels, for a given set of $(K_1,K_2)$ values when compared to the optimal calibration parameters. We note a wide range of solutions equifinal to within 0.01 pixels.

from accurately visible points in the respective orthoimagery rather than from fixed coordinates. For this purpose, we used orthoimagery with a 2.5 cm resolution and we estimate that the point selection was accurate to at least three pixels (7.5 cm). By using recognizable conjugate points as the bounds of the profiles, any scaling errors in the models will be visible as a difference in profile length between points $x$ and $y$. However, Figure 8 shows no visible difference in profile length but a significant vertical offset of $\approx 12$ m. The exact profile length in the X-Y plane was 132.36 m for A1 and 132.79 m for A2. If we consider our estimated measurement error, we obtain a scale discrepancy of 0.26% to 0.38%. This is slightly smaller but consistent with the results in Table IV, which shows scaling ranging from 99.7% to 100.9%. However, closer examination does reveal errors in the vertical scaling of the model. Figure 9 plots the full set of deformations (i.e. residuals) versus elevation. Figure 9 shows a strong correlation between errors and elevations with large scatter capable of enveloping the profiles shown in Figure 7. Closer examinations of the data confirmed this and showed that the vertical scale was either compressed (A2) or dilated (A1). Therefore, focal length calibration remains crucial. In this regard, a re-examination of Table III shows that imagery acquired at NADIR viewing angles (A1, A2, B1 and B2) leads to poor focal length calibrations whilst focal lengths calibrated from convergent view acquisitions (A3, A4, B3 and B4) are all consistent with the initial flat wall experiment.

Overall, the results for rotation, scale, and random error are very encouraging. In terms of rotation, $R_y$ is a rotation around the vertical axis $Z$ and therefore leads to errors in the azimuths of the model; $R_\theta$ is rotation around the X-axis (north and south tilt, after application of $R_y$). This therefore represents a rotation away from verticality and should be considered as an important source of error for surface process science applications. Finally, $R_\phi$ gives the rotation around the X-axis (east and west tilt) after the application of $R_y$ and $R_\theta$. Values of $R_\theta$ are encouraging, most values are below 1° and as low as $-0.12^\circ$. As a reference, our maximum and minimum $R_\theta$ errors of $2.42^\circ$ and $-1.42^\circ$ would translate as vertical errors of 4.23 m and $-2.48$ m (respectively) at a distance of 100 m from the rotation centre. Errors in the scaling parameter are also a potentially important since they will affect any measurements of distance and volume. The results are again encouraging with several values being within $\pm 1\%$ of the correct scale (i.e. 99–101%). However, some results are poor with the minimum value of 94.9% for B3 and a maximum value 104.8% for B4. Random error ($\eta_p$) values are also encouraging. Given the non-uniform vegetation present at Site A, we would expect a precision of 0.3 to 0.4 m when comparing the photogrammetric point clouds to the dGPS bare earth surface. However, at Site B the short-cropped lawn should have little impact on the data. This is consistent with the data in Table IV where random errors are in the area of 0.4 m for Site A and as low as 0.06 m for Site B. This promising result of 0.06, at a flying height of 60 m, would indicate an optimal precision of 0.1% (1:1000) of flying height.

adjustments (James, personal communication). This means that both our work and that of James and Robson (2012) are in essence using a purely image-based SfM process to derive a 3D point cloud in arbitrary space and then using a network of GPS coordinates to calculate a seven-parameter Helmert transform in the least-squares sense in order to georeference this point cloud to map coordinates. The fact that both works reached similar precisions of 1:1000 suggests that the SfM image-based approach is inherently limited to such precisions levels. This observation requires further investigation but should serve as a caution to researchers that view SfM as a method which bypasses the laws of photogrammetry.

Our findings further confirm the recommendations of James and Robson (2014) stating that convergent viewing angles optimize the results of camera self-calibration. However, in addition to the knowledge that convergent viewing angles are qualitatively ‘good’ for camera calibration, our results provide a quantitative approach to calibration assessment. We recommend the following steps (summarized in Figure 10): The camera should be calibrated before aerial acquisition, 20–30 images at convergent angles should be acquired of a flat wall on a recently constructed surface with a good level of texture. In our experiments, we acquired heavily overlapping (>80%) imagery with the optical axis of the camera intersecting the wall surface at an angle of ≈±45°. Rather than implementing the computationally onerous parameter space exploration we

Table III. Self-calibration outputs for all flight experiments.

<table>
<thead>
<tr>
<th>Flight</th>
<th>F [pix]</th>
<th>$K_1 \times 10^{-6}$</th>
<th>$K_2 \times 10^{-15}$</th>
<th>$K_3 \times 10^{-21}$</th>
<th>$K_2' \times 10^{-15}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Site A</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>A1: NADIR 1</td>
<td>2539</td>
<td>-2.48</td>
<td>3.86</td>
<td>-9.48</td>
<td>1.69</td>
</tr>
<tr>
<td>A3: Oblique 1</td>
<td>2317</td>
<td>-2.52</td>
<td>3.84</td>
<td>-9.69</td>
<td>0.04</td>
</tr>
<tr>
<td>A4: Oblique 2</td>
<td>2320</td>
<td>-2.52</td>
<td>3.86</td>
<td>-10.6</td>
<td>0.06</td>
</tr>
<tr>
<td>Site B</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>B1: Nadir 1</td>
<td>2304</td>
<td>-2.45</td>
<td>4.11</td>
<td>-11.9</td>
<td>0.45</td>
</tr>
<tr>
<td>B2: Nadir 2</td>
<td>2229</td>
<td>-2.58</td>
<td>4.12</td>
<td>-13.0</td>
<td>0.02</td>
</tr>
<tr>
<td>B3: Oblique 1</td>
<td>2323</td>
<td>-2.59</td>
<td>4.13</td>
<td>-13.2</td>
<td>0.18</td>
</tr>
<tr>
<td>B4: Oblique 2</td>
<td>2323</td>
<td>-2.58</td>
<td>4.10</td>
<td>-12.2</td>
<td>0.16</td>
</tr>
</tbody>
</table>

Note: $K_2'$ gives the error in $K_2$ when calibrated values of $K_2$ are compared to those predicted by Equation (5).
conducted in Photoscan (leading to Figure 6), we recommend that users simply self-calibrate the convergent imagery in order to obtain an assumed optimal calibration. As a check of this assumption, the flat wall surface point cloud should be regressed in a second-order polynomial model. The second-order terms should be negligible. Once an optimal calibration value is determined, users can replace the Photoscan parameter exploration with the low-impact Monte Carlo approach described here (leading to Figure 5) to determine equifinal solutions in the $K_1-K_2$ (potentially adding $K_3$) parameter space.

Our findings indicate that solutions need to be similar to within a maximum displacement of two pixels. However, if we consider Figures 5 and 6, it becomes apparent that the acceptable domain of equifinal solutions will scale with flying height. In the case of the flat wall experiment where the effective flying height was 10 m (roughly 20% of the width of the wall), equifinality was observed over a wide range of solutions with displacements similar to within 20 pixels. In the case of real data, the flying height was 60 m (50% of the site width) and in this case, we found that displacements similar to two pixels were required. Users operating at lower altitudes can therefore expect satisfactory results from calibrations delivering lens distortions that match optimum values to displacements somewhere between two and 20 pixels. With this target in mind, it is recommended that at least 10% of an image dataset be acquired at angles of 20° to 45° off-NADIR. Once real aerial survey data are processed, the resulting self-calibration can then be checked against the range of equifinal $K_n$ parameters.

Furthermore, the focal length calibration should match to within 0.1% of the initially calibrated value. Since SfM-photogrammetry packages self-calibrate the camera during the initial camera alignment stage, the calibration can be checked on a field laptop without the need for the more computationally demanding stage of dense topography reconstruction. If calibration results are not satisfactory, we

![Figure 7](image_url) Linear profiles from $\alpha$ to $\beta$ with self-calibrated and adjusted calibration values. Deformation amplitude is calculated by comparing sites A1–A3 to the optimal A4 and B1–B3 to the optimal B4. We note that whilst doming is eliminated, the adjustment results in increased surface noise.

![Figure 8](image_url) Effects of focal length calibration errors on datum along $\alpha-\gamma$. Note that profile length in $XY$ is not affected.

Table IV. Error parameters for all flight experiments.

<table>
<thead>
<tr>
<th>Flight</th>
<th>Tx [m]</th>
<th>Ty [m]</th>
<th>Tz [m]</th>
<th>R$\phi$ [°]</th>
<th>R$\theta$ [°]</th>
<th>R$\psi$ [°]</th>
<th>$S$ [%]</th>
<th>$\eta_p$ [m]</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Experiment 1: Site A</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>A1: NADIR 1</td>
<td>4.02</td>
<td>1.32</td>
<td>-2.48</td>
<td>-0.93</td>
<td>-0.38</td>
<td>0.40</td>
<td>99.7</td>
<td>0.46</td>
</tr>
<tr>
<td>A2: NADIR 2</td>
<td>3.31</td>
<td>-1.79</td>
<td>8.88</td>
<td>-0.60</td>
<td>-1.42</td>
<td>-0.34</td>
<td>100.9</td>
<td>0.36</td>
</tr>
<tr>
<td>A3: Oblique 1</td>
<td>0.65</td>
<td>0.35</td>
<td>3.49</td>
<td>-1.34</td>
<td>-0.44</td>
<td>-0.02</td>
<td>99.5</td>
<td>0.55</td>
</tr>
<tr>
<td>A4: Oblique 2</td>
<td>3.04</td>
<td>1.63</td>
<td>0.36</td>
<td>-0.62</td>
<td>-0.78</td>
<td>-0.18</td>
<td>100.4</td>
<td>0.32</td>
</tr>
<tr>
<td>registration error [m]</td>
<td>0.05 ± 0.23</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Experiment 1: Site B</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>B1: NADIR 1</td>
<td>0.82</td>
<td>6.38</td>
<td>-5.73</td>
<td>0.86</td>
<td>-1.20</td>
<td>1.11</td>
<td>98.4</td>
<td>0.12</td>
</tr>
<tr>
<td>B2: NADIR 2</td>
<td>-0.46</td>
<td>4.20</td>
<td>5.19</td>
<td>0.09</td>
<td>0.33</td>
<td>-0.06</td>
<td>95.8</td>
<td>0.24</td>
</tr>
<tr>
<td>B3: Oblique 1</td>
<td>0.79</td>
<td>5.85</td>
<td>1.17</td>
<td>-0.08</td>
<td>-1.05</td>
<td>0.66</td>
<td>94.9</td>
<td>0.15</td>
</tr>
<tr>
<td>B4: Oblique 2</td>
<td>-3.31</td>
<td>2.43</td>
<td>-0.49</td>
<td>0.68</td>
<td>0.23</td>
<td>0.02</td>
<td>104.6</td>
<td>0.06</td>
</tr>
<tr>
<td>registration error [m]</td>
<td>0.00 ± 0.03</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Note: Tx, Ty, Tz give translations (systematic offsets), R$\phi$, R$\theta$, R$\psi$ give Euler rotation angles with R$\theta$ being giving off-vertical tilt angles. S is the scaling error with 100% meaning that the model scale is identical to the actual scale. $\eta_p$ is the precision of the final SfM point clouds, calculated as the standard deviation of the quasi-random field $\eta$. For each site, we give the registration error estimate for the cloud alignments as accuracy ± precision where the accuracy is the mean registration error and precision is one standard deviation of the registration error.
recommend adding extra convergent-view images acquired from a higher altitude. At this point in our research, we do not recommend adjusting calibration values to match equifinal \( K_1 \) solutions and/or using a fixed calibration during photogrammetric processing. Figure 7 clearly shows that whilst this improves the accuracy of the model by eliminating doming and dishing, it increases surface noise and thus lowers the precision of the final model. Furthermore, calibrations values fixed to those of the flat wall experiment also produced increased levels of noise with the associated loss of precision. This is consistent with the findings of Fraser (1997), Förstner et al. (2013), Fraser (2013) and Wolf et al. (2014) that recommend on-the-job self-calibration as best-practice for small format, non-metric, photogrammetry. However, we note that our approach requires a certain pattern of lens distortion which is common for a majority of compact cameras, as well as many DSLR zoom lenses. If we examine the Brown–Conrady lens model in Table I and the lens distortion profiles in Figure 4, it can be inferred that a necessary condition for equifinality in the \( K_1 - K_2 \) parameter space is opposite signs for \( K_1 \) and \( K_2 \). In our case, this is what allows for various combinations of \( K_1 \) and \( K_2 \) to compensate each other. In order to assess transferability of these findings, we conducted a survey of a range of cameras and examined their \( K_1 - K_2 \) radial distortion parameters as obtained by the self-calibration algorithm in Photoscan. Table V shows that in all cases, \( K_1 \) and \( K_2 \) have opposite signs. Furthermore, we have conducted Monte Carlo lens simulations similar to the one shown in Figure 5 on all cameras in Table V, in most cases, we find the same pattern of equifinality expressed as a family of solutions \( K_2 = f(K_1) \) where \( f \) is a linear function with a negative slope. The only exception is the Microdrone 2.0 camera which did not show any equifinality. However, this is a miniature camera weighing 18 g, equipped with a rolling shutter and which is in fact a video camera that can produce single frame stills. Therefore, we conclude that such miniaturized devices are not yet ready to be used as quantitative tools. However, the other cameras in Table V do display equifinality and correlation in their \( K_1 - K_2 \) parameter spaces leading us to conclude our findings with respect to camera calibration are readily transferable.

Once it is confirmed that a self-calibrated block adjustment has been achieved within a family of zero distortion solutions, the rigid-body seven-parameter transformation that is employed to transform a point cloud from the arbitrary space of the initial block adjustment to map coordinates, can successfully produce topographic models to a precision of 0.1% of flying height. This translates to decimetric precisions at altitudes less than 100 m above ground level (AGL). These

**Figure 9.** Altitude dependent errors. Here we show the correlation between absolute altitude and error for flights A1 and A2 which had the poorest calibration of focal length.

**Figure 10.** Summarized DG workflow. The two stages of the proposed workflow. The camera calibration stage aims to identify the focal length and the set of equifinal solutions along with their domain of applicability. This information can then be used in the calibration check, potentially carried out in the field, that allows for a reliable topographic model to be produced with a DG.

The accuracy of DG-derived topographic datasets, correcting for datum shifts, scale, and tilt. The main difference with a traditional GCP survey is that the GCPs in a DG workflow do not need to be well distributed throughout the area being mapped. However, this ground-truthing does not necessarily require survey-grade data or equipment. Vertical datum shift and translation errors in XY (Tx, Ty, Tz) could arguably have no impact where a single-epoch DEM is acquired for purposes such as hydrologic or hydraulic modelling. These errors are the easiest to correct as they require only one point with a known elevation. This can be retrieved from freely available global or national datasets. However, for change detection studies, correction of these linear shifts will be critical. In these cases any given point in the survey area which has remained static can be used to correct translation errors. Scaling errors may have an impact on studies of erosion and volumetric change. Their correction requires two points with a precisely known relative distance. Fortunately, a single scale object with accurately known dimensions in the study site is sufficient. Since the scale parameter used in the seven-parameter transform is isotropic, the scale object can be in any orientation and the resulting scale correction is valid for the whole model. The accuracy of this scale correction will be a function of the size of the object. For corrections to within 0.1%, a scale object separated by 1000 pixels is required which can be obtained from larger man-made features. In certain cases, even Google Earth can provide suitable scales. In the case of B4, we used Google Earth to measure the length of a pale cement border along the path and used this measurement in Photoscan as a scale bar. This improved model quality and resulted in a scaling parameter of 100.2% (down from 104.8%). However, readers should be cautious when using Google Earth data for accurate mapping of any sort. Google Inc. acquires data from multiple sources and there is no stated, universal, quality standard. In fact, care should be taken with scale measurements from remotely sensed imagery, often times the measurement error from certain imagery sources can be larger error than the errors from the sUAS GPS. Tilt errors are significant in the context of surface processes where models of sediment and/or water fluxes will be sensitive to gradient errors. These errors can be corrected with two scale objects set at a right angle and levelled to an accurate horizontal pitch. If using projected map coordinates, it is also recommended to use a compass to align the azimuths of these two scale objects to N-S and E-W directions. The scale object could be located anywhere in the study area, including at the unmanned aerial vehicle (UAV) launch point, which would be useful in situations where the area being mapped is inaccessible. Alternatively, secondary elevation datasets with sub-metric vertical precisions might be used (i.e. existing airborne or terrestrial LiDAR). Here we recommend a simple planar detrending operation based on the difference between the DG topography and the reference topographic data. With the growing availability of high-resolution (metric and sub-metric) datasets, correcting for datum shifts, scale and tilt is now relatively straightforward in many parts of the globe.

Table V. Calibrated focal lengths and $K_i$ parameters for a selection of cameras.

<table>
<thead>
<tr>
<th>Camera Type</th>
<th>Focal [pix]</th>
<th>$K_1$ [f]</th>
<th>$K_2$ [f]</th>
<th>$K_3$ [f]</th>
</tr>
</thead>
<tbody>
<tr>
<td>Canon EOS Rebel T2i (18–135 mm)</td>
<td>4411.34</td>
<td>−0.169</td>
<td>0.144</td>
<td>0.026</td>
</tr>
<tr>
<td>Canon PowerShot A3300 IS</td>
<td>3827.51</td>
<td>−0.040</td>
<td>0.045</td>
<td>−0.013</td>
</tr>
<tr>
<td>Canon PowerShot A4400 IS b</td>
<td>3801.63</td>
<td>−0.082</td>
<td>0.070</td>
<td>−0.003</td>
</tr>
<tr>
<td>Canon PowerShot S10</td>
<td>2930.36</td>
<td>−0.04</td>
<td>0.01</td>
<td>−0.01</td>
</tr>
<tr>
<td>Canon PowerShot SD1000 (IXUS 70)</td>
<td>1822.4</td>
<td>−0.17</td>
<td>0.25</td>
<td>−0.07</td>
</tr>
<tr>
<td>Canon PowerShot SX170 IS b</td>
<td>3823.95</td>
<td>−0.091</td>
<td>0.432</td>
<td>−0.877</td>
</tr>
<tr>
<td>LG Nexus 5 b</td>
<td>3019.65</td>
<td>−0.215</td>
<td>0.385</td>
<td>−0.078</td>
</tr>
<tr>
<td>Microdrone 2.0 HD</td>
<td>1466.43</td>
<td>0.20</td>
<td>−1.40</td>
<td>3.00</td>
</tr>
<tr>
<td>Nikon D1X 35 mm lens</td>
<td>4535.33</td>
<td>−0.10</td>
<td>0.16</td>
<td>−0.12</td>
</tr>
<tr>
<td>Nikon D5200 (18–55 mm)</td>
<td>4635.72</td>
<td>−0.102</td>
<td>0.065</td>
<td>−0.079</td>
</tr>
<tr>
<td>Zoom @ 18 mm</td>
<td>5231.44</td>
<td>−0.15</td>
<td>0.13</td>
<td>0.03</td>
</tr>
<tr>
<td>Sony Nex7, 20 mm lens</td>
<td>1681.51</td>
<td>−0.227</td>
<td>0.055</td>
<td>−0.0056</td>
</tr>
</tbody>
</table>

Note: Most of the calibration data was extracted from real surveys processed with Photoscan. Four of the cameras were calibrated with the camera calibration operations in the OpenCV software libraries. All of the cameras tested have opposite signed $K_1$ and $K_2$ values.

(a)Traditional airborne survey pattern, acquired at NADIR with two altitudes.

(b)Parameters derived from the camera calibration operations in OpenCV (opencv.org).

An examination of Table IV clearly shows that the optimal precision ($\eta_p$) does not coincide with optimal scale and tilt. Corrective steps must be taken to get the best results and in fact, we find that DG from consumer-grade sUAS still requires minimal ground-truthing for most science applications. If future research can improve the quality of DG-derived topography, the DG_workflow might out-compete laser scanning for certain applications with much smaller operational and capital costs. We therefore argue that the combination of low-cost drones, low-cost SFM-photogrammetry and a DG workflow will transform mapping by allowing both specialists and non-specialists to generate topography and 3D virtual landscapes with meaningful levels of accuracy and precision.

We will now proceed to a closer examination of the errors and likely sources that affect the DG approach with sUAS. If survey-grade RTK-GPS is available, the workflow presented here still has implications. Traditionally, it is recommended to have in excess of 20 GCPs well distributed in the XYZ space of the study site (Carbonneau et al., 2003). However, in combination with a DG workflow, three to four survey-grade GCPs, preferably co-located with pseudo-invariant features, could accurately correct for datum shifts, scale, and tilt. The main difference with a traditional GCP survey is that the GCPs in a DG workflow do not need to be well distributed throughout the site. For example, in a traditional photogrammetry survey, the network of GCPs must not be co-planar or co-linear. However, in the context of a DG workflow, co-planar or linear GCPs can be tolerated and used to correct tilt and scale (respectively) because the actual 3D model is first solved in...
an arbitrary space without any position information. Once this is done, the available position information is used in a least-squares sense to calculate the seven-parameter Helmert transform that will georeference the point cloud to map coordinates (James and Robson, 2012; Fonstad et al., 2013). This could allow for topographic surveys in inaccessible areas provided that the edges of these sites are accessible. However, if limited but high-accuracy GCPs are used, readers should note that these GCPs will have a dominant, but not absolute, influence on the final position and orientation of the point cloud. In an auxiliary experiment, we used an additional dataset acquired with another P3P on the White River, Vermont, USA. This data was comprised of a total of 220 images and 11 GCPs surveyed with RTK-GPS. First, we produced a point cloud with the DG workflow described here. Second, we produced a point cloud with only three GCPs in one corner of the site and third, we produced a point cloud with both the DG information and the same three GCPs. When we compare the predicted position of the remaining nine GCPs to their actual surveyed location, we find that the DG only case has an offset magnitude of 1.190 m, the three GCPs case an offset magnitude of 0.022 m and the DG with three GCPs case an offset magnitude of 0.006 m. Clearly, the addition of limited ground control has had a dominant effect on the georeferencing quality. However, the camera positions used in the DG workflow have still contributed to reducing the error. This is likely caused by the large number of cameras still having an impact on the least-squares solution which determines the seven-parameter Helmert transform. Therefore, it can be concluded that a sparse network of GCPs can be supported by a dense network of camera positions. This observation provides a way forward for change-detection studies where multiple static points in the study site will be needed to correct for scale and datum shifts. Furthermore, the error tolerance of each specific study should be used to establish the sUAS flying height as one thousand times the maximum error.

The residual error of 0.1% of flying height (here 0.06 m with flights at an altitude of 60 m AGL) can at best be qualified as a ‘good’ performance. With the use of GCPs, significantly better sUAS photogrammetry precisions have been reported (Eltner et al., 2015; Woodget et al., 2015) and therefore the next steps should be the understanding of DG non-affine errors. However, it is difficult to assess this explicitly. Proprietary SIM-photogrammetry software packages such as Photoscan Pro protect many details in their matching processes for commercial reasons. At the moment, carefully designed empirical experiments appear to be the only way forward and further research on the relationship between surface noise amplitude and pattern matching performance in the DG workflow will be required to achieve survey-grade outputs. Additionally, improving the positional accuracy of each image geotag should be a priority. Bláha et al. (2012) have demonstrated that sub-metric positioning of sUAS can be achieved with on-board RTK-GPS positioning. Furthermore, Chiang et al. (2012) and Milik and Gabrilik (2015) have demonstrated that when RTK-GPS positioning is propagated through a photogrammetric solution, the position of GCPs can also be predicted with decimetric accuracies. Moreover, Eling et al. (2015) report predictions of GCP positions with centimetric accuracy. Whilst these authors use experimental RTK-GPS equipment and high-cost platforms, low-cost RTK-GPS units are appearing. SwiftAv inc. is currently marketing their PIKSI GPS, which is a fully miniaturized GPS capable of real-time kinematic differential corrections. At a cost of $US$900, the units are well suited to low-cost sUAS. Furthermore, RTK-enabled UAVs are appearing on the consumer market with models aimed at both the scientific (e.g. the DJI Matrice 600 hexacopter) sectors. These units will inevitably make their way into the consumer market and we expect publications of results from these RTK-equipped platforms in the near future strengthen the case for the use of DG in sUAS photogrammetry. However, we note that our findings re-emphasize the importance of camera calibration as a crucial factor in accurate topography reconstruction. The addition of RTK-GPS to sUAS might provide considerable improvements to the performance of the DG workflow, but it will not obviate the need for accurate camera calibration. This point emphasizes the fact that SIM-photogrammetry is in no way exempt from the principles of photogrammetry. Advances in image matching algorithms from the area of computer vision, integrated into the photogrammetric workflow, have clearly enabled a step-change in image-based topography generation. However, practitioners of SIM-photogrammetry still require traditional, ‘pre-SIM’ knowledge of photogrammetry. The illusion that SIM is not photogrammetry must now be dispelled.

Conclusion

The problem of producing topography from DG sUAS imagery poses many significant challenges and opportunities. When we compare our results to those published by other authors, it is clear that the DG workflow results in the expected higher levels of error. However, our results and innovative approach to error characterization indicate that current consumer-grade drones along with low cost SIM-photogrammetry packages and a DG workflow can produce topographic data with sufficient quality for a limited number of applications. Given that the cost of RTK-GPS equipment is generally above ~£10 000 and that the low-cost drones used here are in the area of £1000 to £2000, our DG approach offers a reduction of costs of a full order of magnitude which is further reduced by the lower field effort and associated person-hours required by DG. We therefore argue that further development and integration of DG into the UAS/SIM workflow has significant implications for topographic survey that justify further research and development. With the cost of consumer-grade drones still declining, the DG approach could allow for high quality, meso-scale topography to be produced at dramatically reduced costs. This mass-production of topographic data and associated sub-metric resolution imagery will have a transformative impact on all Earth surface process sciences as well as the topographic survey and natural disaster management industries.

Acknowledgments—The authors would like to thank Drs Matt Westoby and Mike James for very constructive reviews. The authors would also like to thank Howard Harper and Max Densmore for piloting assistance during data acquisition in the UK.

References


