

Camera orientation, calibration and inverse perspective with uncertainties: A Bayesian method applied to area estimation from diverse photographs

Grégoire Guillet^{a,*}, Thomas Guillet^b, Ludovic Ravanel^a

^a Université Grenoble-Alpes, Université Savoie Mont-Blanc, Laboratoire EDYTEM, 5 boulevard de la mer Caspienne, 73370 Le Bourget-du-Lac, France

^b Department of Physics and Astronomy, Stocker Road, University of Exeter, EX4 4QL, United Kingdom



ARTICLE INFO

Keywords:

Inverse perspective
Spatial resection
Camera calibration
Bayesian methods
Lens distortion
Digital elevation models

ABSTRACT

Large collections of images have become readily available through modern digital catalogs, from sources as diverse as historical photographs, aerial surveys, or user-contributed pictures. Exploiting the quantitative information present in such wide-ranging collections can greatly benefit studies that follow the evolution of landscape features over decades, such as measuring areas of glaciers to study their shrinking under climate change. However, many available images were taken with low-quality lenses and unknown camera parameters. Useful quantitative data may still be extracted, but it becomes important to both account for imperfect optics, and estimate the uncertainty of the derived quantities. In this paper, we present a method to address both these goals, and apply it to the estimation of the area of a landscape feature traced as a polygon on the image of interest. The technique is based on a Bayesian formulation of the camera calibration problem. First, the probability density function (PDF) of the unknown camera parameters is determined for the image, based on matches between 2D (image) and 3D (world) points together with any available prior information. In a second step, the posterior distribution of the feature area of interest is derived from the PDF of camera parameters. In this step, we also model systematic errors arising in the polygon tracing process, as well as uncertainties in the digital elevation model. The resulting area PDF therefore accounts for most sources of uncertainty. We present validation experiments, and show that the model produces accurate and consistent results. We also demonstrate that in some cases, accounting for optical lens distortions is crucial for accurate area determination with consumer-grade lenses. The technique can be applied to many other types of quantitative features to be extracted from photographs when careful error estimation is important.

1. Introduction

A large amount of quantitative physical landscape information can be extracted from terrestrial, aerial and satellite imagery using various photogrammetric techniques (Streilein, 1994; Gruen and Li, 1995; Haala and Brenner, 1999; Küng et al., 2012; Feurer and Vinatier, 2018). Inverse perspective methods (e.g. monoplotting), as reviewed by Criminisi, 2001; Förstner and Wrobel, 2016, aim at extracting referenced spatial data from a single picture. Such methods have been used to extract data from either aerial, satellite or terrestrial imagery (Jordan et al., 2005; Bozzini et al., 2012; Murtyoso et al., 2014; Produit et al., 2016). Inverse perspective methods are particularly used in the study of Earth surface processes and landscape evolution to produce or update geological and geomorphological map data (Warner et al., 1993; Jauregui et al., 2002; Micheletti et al., 2015; Scapozza et al., 2016) or, among others, in civil engineering and building stability assessment

(Murtyoso et al., 2014). The methods have also found a particular echo in the community of cryospheric sciences, as they allow to reconstruct and monitor the evolution of glaciers over different time scales, ranging from centennial to annual fluctuations (Wiesmann et al., 2012; Piermattei et al., 2015; Čekada et al., 2016), as well as further the understanding of glacier mass balance processes (Chapuis et al., 2010).

Inverse perspective methods also allow to tap into a wealth of quantitative information present in large and diverse databases of images readily accessible from the Internet, such as historical records, aerial surveys, or user-contributed pictures. However, these images are of uneven quality: many were taken without scientific intent, often with low-quality lenses and unknown camera parameters, and are sometimes only available in low resolution. These limitations can introduce significant uncertainties and biases in the information obtained from camera orientation and calibration. Useful quantitative data may still be extracted, but accounting for potential lens distortions and

* Corresponding author.

E-mail address: gregoire.guillet@univ-smb.fr (G. Guillet).

quantifying the uncertainty of the results become important.

In this paper, we present a method to address both goals, applying it to the estimation of the area of landscape features, with the determination of the areas of mountain glaciers in mind. The present technique is a two-step process based on Bayesian inference.

First, the unknown camera parameters, including lens optical distortions, are estimated using a Bayesian formulation of the camera orientation with calibration problem, in the form of a spatial resection problem: the posterior probability density function (PDF) of camera parameters is obtained from matches between 2D points in the image and 3D points in world coordinates, together with any available prior information. Bayesian approaches to camera calibration were presented by several authors, either based on finding a single value of the parameters which maximize a posterior distribution (e.g. Valkenburg, 1998; Zhang, 2000) or using the whole resulting posterior distribution more extensively (e.g. Sundareswara and Schrater, 2005). In our work, we keep the full statistical information contained in the posterior distribution of camera parameters, by generating samples distributed according to the posterior distribution using Markov chain Monte Carlo (MCMC) sampling.

In a second step, the posterior PDF of the feature area is derived from the posterior of camera parameters by solving an inverse perspective problem. The outline of the feature of interest is manually traced on the photograph as a polygon, which is then back projected from the 2D image onto the 3D world using a digital elevation model (DEM). This back projection step accounts for uncertainties on the camera parameters, and attempts to model possible systematic errors introduced by the polygon tracing step, together with uncertainties introduced by the DEM. In particular, we propose a model for DEM errors for which both the root mean square error (RMSE) and spatial autocorrelation scale are locally varying. The resulting PDF of the back projected 3D feature area therefore contains information about most of the uncertainties of the process.

More generally, we attempt to unify the camera orientation with calibration, uncertainty modeling and inverse perspective problems into a statistically consistent framework which can be extended to similar classes of problems and uncertainty models.

We stress that given the diverse nature and sources of our target images and the fact that many were taken using low-quality or unknown equipment, the focus of this paper is more on uncertainty estimation than on very accurate photogrammetric techniques. The reconstructed camera location, for example, cannot be expected to be more accurate than a few meters, given that we will be working with low resolution landscape images, and that our 3D ground control point coordinates will not come from precision geodetic sources.

We start by describing our Bayesian formulation of the camera orientation with calibration problem in Section 2, and our polygon back projection method in Section 3. Section 4 describes details of our implementation, including and the posterior probability density sampling process. In Section 5, we present test problems for validation, before discussing the results in Section 6 and concluding in Section 7.

2. Bayesian formulation of camera orientation and calibration

2.1. Camera orientation and calibration

Extracting metric measurements from digital images requires estimating the parameters of the imaging camera used to take the picture, such as its position and orientation, focal length, and possibly other optical properties. For typical orientation problems, this may be done by matching points with known 3D world coordinates with their corresponding 2D projections in the image under study. Camera orientation and calibration then consists in finding the camera parameters that best reproduce the 2D points by projecting their matching 3D point using the chosen camera model. The process is illustrated in Fig. 1. The control points are typically selected as characteristic features on the 2D

image (such as buildings, peaks, road crossings, etc.) for which the corresponding points in 3D world coordinates can readily be identified. The 3D coordinates of the control points may be determined using a Digital Elevation Model (DEM) (Jauregui et al., 2002; Willneff et al., 2005; Stockdale et al., 2015) or other sources of cartographic and elevation data.

In the present paper, we note θ the vector of camera parameters; those will be further described in Section 2.3 as we discuss the camera model, and are summarized in Section 2.3.3 for convenience. The camera calibration problem may be formulated as finding the camera parameters θ satisfying:

$$\mathbf{d}_{2D} \simeq \mathbf{C}(\mathbf{d}_{3D}; \theta), \quad (1)$$

where θ belongs to a set of admissible camera parameters, \mathbf{d}_{2D} and \mathbf{d}_{3D} are the set of 2D and 3D points respectively, and \mathbf{C} is the camera projection operator, representing the photographic projection operation, and \simeq represents some notion of closeness for points in the image plane.

The actual process of determining θ may depend on the nature of the camera model described by \mathbf{C} . For ideal rectilinear camera models, θ is generally computed via the so-called collinearity equations (Ghosh, 2005; Sheng, 2005), which may be solved by least squares methods (Gruen, 1985; Heipke, 1992; Strausz, 2001; Bozzini et al., 2012). More generally, the process can be seen as a nonlinear minimization of the reprojection errors (i.e. the discrepancy between the 2D projections $\mathbf{C}(\mathbf{d}_{3D}; \theta)$ of the 3D points projected using θ , and the corresponding 2D control point locations \mathbf{d}_{2D}), which is also applicable in the presence of optical lens distortions in the model.

Beyond minimization of reprojection errors, camera calibration can also be formulated probabilistically within a Bayesian framework (e.g. Valkenburg, 1998; Sundareswara and Schrater, 2005). In this approach, a posterior distribution for θ is estimated, and may then be used to infer a suitable single value for θ , for example using the mode (maximum) of the posterior distribution (Valkenburg, 1998), or by drawing random samples from it which can be used to construct other point estimates (Sundareswara and Schrater, 2005).

In this work, we use an approach similar to Sundareswara and Schrater (2005), and rely on a Bayesian statistical framework to characterize a distribution of values of θ , compatible with a set of given 2D/3D matches. Instead of producing a single point estimate for θ , we obtain a probability density function on the camera parameters. It is conditional on the observed matches, and any prior information that we have on probable values of θ .

2.2. Bayesian formulation

Consider the camera calibration problem for a single photograph, for which we have a set of n matches between 3D points $\mathbf{d}_{3D} = \{(x_p, y_p, z_p), p = 1 \dots n\}$, and 2D projected points on the image $\mathbf{d}_{2D} = \{(u_p, v_p), p = 1 \dots n\}$. Collectively, we note $D = (\mathbf{d}_{3D}, \mathbf{d}_{2D})$ this data available to reconstruct the camera parameters θ . We further note I the set of all prior information available about the picture which is known before considering D , such as possible camera orientations or any information known about the lens. I also includes the assumptions about the projection model itself. The Bayesian camera calibration problem amounts to finding $p(\theta|D, I)$, which is the probability density of θ conditional to knowing both D and I , also known as the *posterior PDF* of θ . Applying Bayes' theorem to our problem, we can write:

$$p(\theta|D, I) = \frac{p(D|\theta, I) p(\theta|I)}{p(D|I)}. \quad (2)$$

The terms appearing on the right hand side play distinct roles. $p(D|\theta, I)$, the *likelihood*, is the probability density of observing the matches described by D and the model (1) if we assume the camera parameters θ and prior information I to be known. In our model, this term is computed by evaluating the 2D projected locations of the 3D

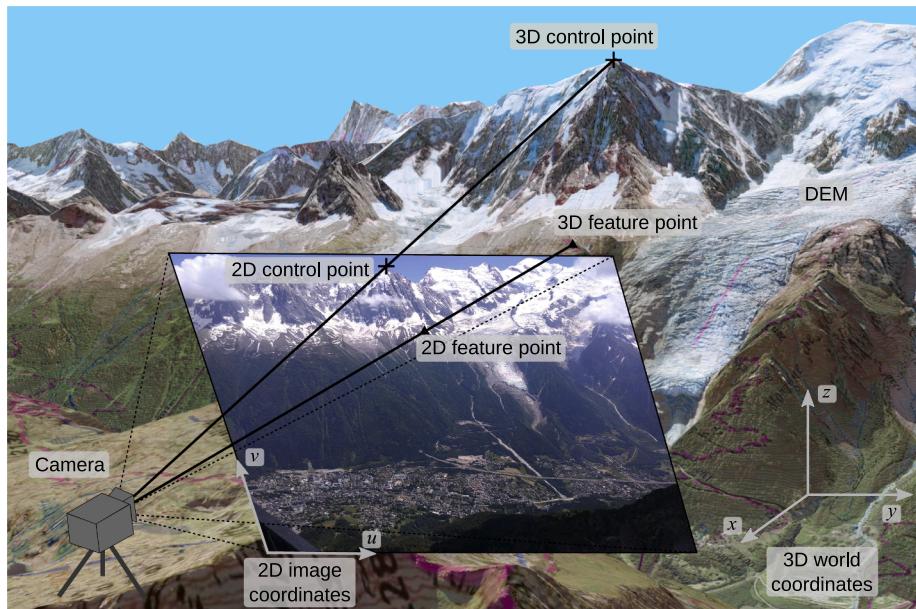


Fig. 1. Overview of the setup for camera orientation and inverse perspective problems. Pairs of matching 2D and 3D control points are used to determine the location of the camera, and establish a mapping between 2D image plane coordinates and 3D world coordinates using a digital elevation model (DEM).

points using our camera model \mathbf{C} , and providing a statistical interpretation of the 2D reprojection errors. In Section 2.3, we detail our camera projection model for \mathbf{C} , and cover the likelihood itself in 2.5.

$p(\theta|I)$ is the *prior*, and encodes all *a priori* information assumed about the image, which might be gathered from embedded Exif (exchangeable image file format) picture metadata or just from knowledge of the studied area. We discuss the prior term in detail in Section 2.4.

Note finally that $p(D|I)$ is independent of θ , and can be recovered by requiring that $\int p(\theta|D, I)d\theta = 1$. The value of this constant is of no practical significance in the rest of this work, so we shall omit it and remember that the posterior (2) is defined up to a constant multiplicative factor.

Given D and I , together with prescriptions for the prior and likelihood terms, Eq. (2) defines a function from the multi-dimensional space of parameters θ into real numbers, which completely characterizes the distribution of θ .

2.3. Camera model

The camera model \mathbf{C} describes how 3D world points (x, y, z) are projected into 2D image points (u, v) , given camera parameters θ :

$$(u, v) = \mathbf{C}(x, y, z; \theta). \quad (3)$$

The camera model is at the core of the likelihood term $p(D|\theta, I)$ appearing in Eq. (2): given camera parameters θ and prior information I , the camera model is used to project 3D control points onto the 2D image plane. The resulting reprojection errors between the resulting and actual 2D control point coordinates will be used to provide a statistical prescription for $p(D|\theta, I)$, as is described in Section 2.5.

The role of \mathbf{C} is to emulate the behavior of a realistic camera. Our camera model first performs a simple rectilinear perspective projection based on a pinhole camera model. However, real optics often exhibit deviations from the ideal pinhole camera model due to optical lens distortions, and the resulting errors can be significant for consumer-quality lenses. Because our set of images includes photographs captured using unknown and presumably low-quality optics, we also include a model for lens distortion in order to minimize parameter bias due to distortion errors. The resulting combined model of a pinhole camera with radial lens distortions is considered a good approximation of most lenses (Heikkila and Silven, 1997; Kannala and Brandt, 2006).

2.3.1. Pinhole camera model

The pinhole camera model is commonly used in computer vision, representing a camera with infinitesimally small aperture and perfect rectilinear projection (Sturm et al., 2011; Sturm, 2014). In our implementation, it is parameterized by the location (x_c, y_c, z_c) of the camera optical center, three Euler angles (α, β, γ) representing the orientation of the camera in terms of yaw, pitch and roll, the camera focal length f , and the location (c_u, c_v) of the lens principal point in the image plane. After the pinhole camera projection step, we obtain points (u, v) in the image plane, to which we apply the lens distortion model.

Note that we implemented a rotation parametrization for the Euler angles for oblique terrestrial or aerial photographs, in order to avoid the typical problems with Euler angles close to their degenerate configurations. We find that although they suffer from degeneracies, Euler angles are more convenient to work with than other representations of orientations such as quaternions, because they allow specification of the camera orientation in terms of yaw, pitch and roll, which is intuitive and lends itself well to specifying priors and controlling results.

2.3.2. Lens distortion model

Our implementation of lens distortion relies on a polynomial expansion of radial distortion around the lens principal point, following e.g. (Brown, 1971; Weng et al., 1992; Wang et al., 2008). We adopt in particular the exact model and notations of the PTLens correction model,¹ as it allows using the compatible Lensfun² database of distortion coefficients to derive a distortion prior for consumer optics (see Section 2.4.5).

Radial distortions have been found to provide a very good model for many non-ideal lenses, while avoiding ill-posed parameter determination problems present in more general models (see e.g. (Zhang, 2000) and references therein for a more detailed discussion). Given a pixel (u, v) obtained from the pinhole camera projection, on an image of size (w, h) pixels, we first compute the central normalized coordinates (\tilde{u}, \tilde{v}) :

$$\tilde{u} = (u - wc_u)/L, \quad (4)$$

¹ https://wiki.panotools.org/Lens_correction_model.

² <https://lensfun.github.io/>.

$$\tilde{v} = (v - hc_v)/L, \quad (5)$$

where L is a scale factor in pixel units, and $(c_u, c_v) \in [0, 1]^2$ define the location of the lens principal point (i.e. the intersection of the optical axis with the image plane), with the geometric center of the image corresponding to $(c_u, c_v) = (0.5, 0.5)$. In this work, we treat c_u and c_v as unknown parameters in the camera model. The scale factor is chosen as prescribed in the PanoTools distortion model, and we take $L = \min(w, h)/2$.

The radial distortion model transforms the point (\tilde{u}, \tilde{v}) of an ideal rectilinear lens image, located at normalized radius $\tilde{r} = \sqrt{\tilde{u}^2 + \tilde{v}^2}$ from the optical axis, into a point (\tilde{u}', \tilde{v}') of a distorted image according to:

$$\tilde{u}' = g(\tilde{r})\tilde{u}, \quad (6)$$

$$\tilde{v}' = g(\tilde{r})\tilde{v}\eta^{-1}, \quad (7)$$

where $g(\tilde{r})$ is a polynomial in \tilde{r} controlling the lens distortion, and $\eta > 0$ accounts for possible non-uniform rescaling of the image: $\eta = 1$ corresponds to images with perfect 1:1 pixel aspect ratio, whereas $\eta > 1$ and $\eta < 1$ correspond to images which have been stretched horizontally or vertically respectively. For g , we follow PanoTools and take:

$$g(\tilde{r}) = d_a\tilde{r}^3 + d_b\tilde{r}^2 + d_c\tilde{r} + d_d, \quad (8)$$

where d_a, d_b, d_c, d_d are real coefficients parameterizing the lens distortion. To anchor the global scaling of the distortion coefficients and break the degeneracy with the focal length f , we fix the scale at radius $\tilde{r} = 1$ by setting $g(1) = 1$, which yields $d_d = 1 - (d_a + d_b + d_c)$. The actual lens distortion parameters appearing in θ are therefore (d_a, d_b, d_c) , for which $d_a = d_b = d_c = 0$ corresponds to an ideal rectilinear lens.

Whenever we need to reconstruct scene rays from image plane pixels, for example when back projecting the polygon of interest from the image plane to the 3D world, it is necessary to invert the transformation (6), (7). We perform this inversion numerically using Newton–Raphson iterations.

2.3.3. Summary of camera model parameters

The camera model described in Section 2.3 defines the whole vector of camera parameters θ :

$$\theta := (x_c, y_c, z_c, \alpha, \beta, \gamma, f, c_u, c_v, d_a, d_b, d_c, \eta), \quad (9)$$

where

$[x_c, y_c, z_c]$ is the camera location in 3D world coordinates, $[\alpha, \beta, \gamma]$ are the camera orientation Euler angles (yaw, pitch, roll) in world coordinates, with respect to some rotation parametrization, $[f]$ is the camera focal length, $[c_u, c_v]$ describe the location of the lens principal point, the geometric center of the image being at $(c_u, c_v) = (0.5, 0.5)$, $[d_a, d_b, d_c]$ are the independent radial distortion parameters, $[\eta]$ is the image pixel aspect ratio correction.

2.4. Priors

Now that we have detailed the camera model \mathbf{C} and the corresponding parameters θ , we turn to the description of the prior term $p(\theta|I)$ in Eq. (2), which captures all information about θ that is known regardless of any information coming from the matches D . I may contain information obtainable from the DEM, from Exif metadata in the picture, or from human expertise in analyzing the photograph. From (2) is the joint prior PDF for all the camera parameters:

$$p(\theta|I) = p(x_c, y_c, z_c, \alpha, \beta, \gamma, f, c_u, c_v, d_a, d_b, d_c, \eta|I). \quad (10)$$

Given only prior information, it is reasonable to assume that some of these parameters are independent. To simplify the specification of the prior, we assume that it factorizes as:

$$\begin{aligned} p(\theta|I) &= p(x_c, y_c, z_c|I) \times \\ &p(\alpha|I) p(\beta|I) p(\gamma|I) p(f|I) \times \\ &p(c_u, c_v|I) p(d_a, d_b, d_c|I) p(\eta|I). \end{aligned} \quad (11)$$

This amounts to the user specifying independent prior distributions for the camera position (x_c, y_c, z_c) , the camera attitude angles (α, β, γ) , the camera focal length f , the lens principal point (c_u, c_v) , and finally the distortion parameters (d_a, d_b, d_c) and pixel aspect ratio η .

2.4.1. Camera location

To define a prior on the camera location (x_c, y_c, z_c) , we first recognize that prior information may constrain the altitude z_c depending on the position (x_c, y_c) , and it is therefore useful to write:

$$p(x_c, y_c, z_c|I) = p(z_c|x_c, y_c, I) p(x_c, y_c|I). \quad (12)$$

To define $p(x_c, y_c|I)$, we delimit a polygonal area inside which we take this probability density to be a constant, and outside of which it is zero. The choice of altitude prior $p(z_c|x_c, y_c, I)$ now depends on whether the photograph is taken from the ground surface or not.

For aerial photographs, we assume that the (x_c, y_c) location does not provide additional information on the altitude z_c , and instead we choose a uniform probability density over some range $[z_{\min}, z_{\max}]$ specified by the user, and possibly determined from local geography. Most aerial survey photographs of the Mont Blanc massif of interest to us are, for instance, taken from altitudes between 6000 m and 10000 m.

For terrestrial photographs, we assume that the camera was located close to the ground surface. Given (x_c, y_c) , the local altitude $h(x_c, y_c)$ of ground level may be obtained from the DEM. Because of vertical errors in the DEM in particular in steep terrain, and because the camera may not be exactly at ground level, we take $p(z_c|x_c, y_c, I)$ to be a Gaussian centered on the DEM altitude $h(x_c, y_c)$, with a standard deviation of 5 m, to account for both sources of vertical errors (see 5.1.3 for more information on the DEM used for this study).

2.4.2. Focal length

For $p(f|I)$, we use a uniform prior in $\ln f$, i.e. $p(f|I) \propto 1/f$, truncated to a range of allowed values $f \in [f_{\min}, f_{\max}]$ which are determined from the available prior information.

The choice of a log-uniform prior for f is motivated by the fact that in the pinhole camera model, f plays the role of a multiplicative scale parameter. This prior is scale-invariant: it ensures that, in absence of any additional information on f , the chosen prior is insensitive to the overall normalization of f . This reduces the sensitivity of the inference process to factors such as the sensor size, and makes it possible to easily choose the normalization of f (e.g. to work with 35 mm equivalent focal lengths, see below). Additionally, this prior enforces $f > 0$ regardless of the chosen bounds on $\ln f$. Translation- and scaling-invariant priors are discussed in more detail in e.g. (Sivia and Skilling, 2006). f may be specified in any unit, as long as the bounds $[f_{\min}, f_{\max}]$ and the image plane size (sensor size) in the pinhole camera model equations are scaled accordingly. For the pinhole camera model, we use the focal length in pixel units f_{px} , given by:

$$\frac{f_{\text{px}}}{w} = \frac{f}{W}, \quad (13)$$

where w is width of the image in pixels, and W is the width of the physical image sensor area corresponding to w , expressed in the same units as f .

The prior range $[f_{\min}, f_{\max}]$ is determined from available picture information. The value of w is always available from the picture file under study. For professional and scientific photographs such as aerials, both f and W are usually known to a good accuracy, usually in millimeters. In this case, $[f_{\min}, f_{\max}]$ can be specified according to the precision at which f is known (e.g. due to rounding). With some consumer digital cameras, f may be recorded in the Exif metadata; in case the sensor size W is then readily available, (13) may then be used directly. If not, Exif metadata

may still store the so-called 35 mm film equivalent focal length, e.g. as standardized in [CIPA Standardization Committee \(2005\)](#), which indirectly factors in sensor size information from the camera. In this case, we typically constrain f to be within $\pm 5\text{--}10\%$ of the Exif indicated value, to allow for rounding and potential inaccuracies in the Exif metadata. Finally, whenever no better prior information is available, we rely on visual inspection of the perspective in the photograph to determine a plausible range $[f_{\min}, f_{\max}]$. Here again, this interval may be conveniently specified in terms of 35 mm equivalent focal length, since it provides a good proxy for angular field of view that is also intuitive to photographers, and can then be converted back into pixel units.

2.4.3. Camera angles

$p(\alpha|I)$ is the camera yaw prior, and generally comes from the user's knowledge of the geographical area. For convenience, we choose the reference camera orientation so that the yaw angle is measured with respect to grid North, and therefore corresponds to the camera's viewing azimuth. Depending on features visible on the photograph, it is often possible to constrain the range of possible camera azimuths. We use a uniform prior with user-specified left and right azimuth bounds. Note that we treat all three Euler angles as independent and uniformly distributed³.

$p(\beta|I)$ and $p(\gamma|I)$ encode the independent prior distributions for camera pitch and roll angles. These can usually be estimated visually on most images, with photographs pointing down having negative β , and most terrestrial photographs having a flat horizon and γ close to 0. Here as well, we typically use uniform priors over user-specified ranges.

2.4.4. Lens principal point

To constrain the location of the lens principal point (c_u, c_v) , we assume that the uncertainties in the u and v directions are independent, and therefore $p(c_u, c_v|I) = p(c_u|I)p(c_v|I)$. We further assume the principal point to reside within the image: $(c_u, c_v) \in [0, 1]^2$. A natural way to specify such a prior is to assume independent beta distributions for c_u, c_v .

Consumer cameras and general prior. In absence of more specific prior information, we take $c_u, c_v \sim \text{Beta}(3, 3)$, which imposes only loose constraints on the location of the optical axis, while favoring locations close to the geometric center of the image, at $(c_u, c_v) = (0.5, 0.5)$. This prior, represented in Fig. 2, still allows principal points far away from the image center, and is broad enough to accommodate cropped images.

Aerial photographs with fiducial marks. Professional aerial images, such as the aerials used in this study, feature fiducial marks which can be used to determine the location of the principal point, and therefore are a part of our prior information I .

Taking the horizontal principal point location c_u as an example, we assign a Beta(a, b) prior to c_u , with shape parameters a and b set by imposing the mean and standard deviation of the distribution. The mean of Beta(a, b) is set to match the horizontal principal point location, determined as the center of gravity of the fiducial markers. The standard deviation encodes the *a priori* uncertainty about the relationship between the fiducial marks and the actual principal point. Mapping agencies impose strong compliance constraints on survey cameras, including upper bounds on the distances between the fiducial-indicated principal point, the principal point of autocollimation, and the principal point of symmetry. Typical tolerances are of the order of 30 micrometers ([Tayman, 1984](#); [US Geological Survey, 2012](#)), translating into 1–2 pixels at the resolution of the aerial scans available to us. To account for pixel-level errors in the determination of the fiducial locations, we set the standard deviation of our Beta(a, b) prior to

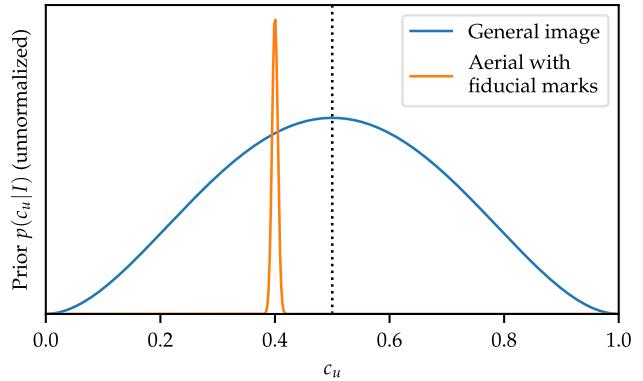


Fig. 2. Illustration of prior distributions for the horizontal location c_u of the lens principal point. The dotted vertical line at $c_u = 0.5$ marks the horizontal center of the image. For general images, a Beta(3, 3) distribution is used (blue curve), which provides a broad prior favoring central locations in the image. Whenever fiducial marks are present, we use a peaked beta distribution centered around the fiducial location of the optical center, with a standard deviation of 5 pixels (illustrated here by the orange curve for a fiducial center point at 400 pixels and an image width of 1000 pixels). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

correspond to 5 image plane pixels. We verify that the midpoints of all segments joining pairs of opposite fiducial marks fall within this distance of the determined principal point.

The prior distribution for c_v in the vertical direction is constructed in an identical way, with its own independently derived a, b parameters. Example priors, without and with fiducial marks, are illustrated in Fig. 2. In practice, we find the principal point to be very well constrained by the control point data with our test images.

2.4.5. Distortion priors

For the distortion parameters d_a, d_b, d_c we use two types of priors depending on the source of the photograph:

Consumer-grade or unknown cameras For photographs assumed to come from general consumer cameras, we derived a joint prior on the three parameters based on the [Lensfun](#)⁴ database of calibrated cameras and lenses. To account for the observed correlations between distortion parameters, we model $p(d_a, d_b, d_c|I)$ as a multivariate Gaussian distribution with mean and covariance matrix estimated from the database entries. The resulting prior distribution is only an approximate match to the data, but we expect it to perform well for reasonably good lenses.

Aerial photography. Aerial photography lenses typically have tightly controlled radial distortions; published tolerances allow of the order of 10 mm in the whole field (e.g. [Tayman, 1984](#); [US Geological Survey, 2012](#)). The French Institut Géographique National (IGN) have graciously provided camera calibration certificates for the aerial missions whose pictures are used in this study; indeed we could verify that the calibrated radial lens distortions never exceed 10 mm. This amount of distortion represents sub-pixel errors for the aerial photographs used in this paper. Nevertheless, we still construct an appropriate prior and fit the distortion coefficients. We constrain d_a, d_b, d_c by imposing a likelihood on the maximum distortion over the whole half-diagonal of the image, modeled as a χ distribution with 1 degree of freedom with some scale s in μm :

$$\max_{0 \leq \tilde{r} \leq \sqrt{2}} |g(\tilde{r})\tilde{r} - \tilde{r}| \sim \chi_1(s/L). \quad (14)$$

Note that this results in strong correlations between the three distortion parameters, as their individual effect can cancel out in Eq. (8). A trivariate distribution was constructed using MCMC sampling of the above

³This simplifying choice is adequate for camera attitudes away from the gimbal lock positions of the chosen Euler angles. In other situations, it would be necessary to define a uniform probability distribution on camera attitudes. This may be done using the Haar measure on the group $SO(3)$ of 3D rotations.

⁴<http://lensfun.sourceforge.net/>.

likelihood; the resulting sampled distribution in d_a, d_b, d_c could then accurately be fitted with a multinormal distribution. The scale $s \approx 3.3 \mu\text{m}$ is determined such that the resulting maximum distortion for this distribution of parameters approximately follows a χ_3 distribution with a 95th percentile of $10 \mu\text{m}$.

2.4.6. Aspect ratio prior

The parameter η appearing in Eq. (7) describes the pixel aspect ratio of the image, to account for possible non uniform rescaling after the image was taken. For many images, we expect η to be very close to 1, and with deviations of no more than 15–20%, as strongly distorted aspect ratios will be detectable by visual inspection. To ensure that η remains positive and obtain a symmetric multiplicative prior around 1, we specify the prior on $\ln\eta$.

For consumer-grade or unknown cameras, we assume $\ln\eta$ to be normally distributed with mean 0 and standard deviation 0.05, resulting in η being close to normally distributed with mean 1 and standard deviation 0.05.

For aerial photography, the aspect ratio can be determined from the horizontal and vertical dimensions of the rectangle delimited by the fiducial marks. We use a Gaussian prior on $\ln\eta$, centered on the value $\ln\bar{\eta}$ determined from the fiducials, and with standard deviation $|\ln\bar{\eta}|$, so that the square aspect ratio $\ln\eta = 0$ is always allowed by the prior. In practice, we always find $\ln\bar{\eta}$ to be very close to 0 for aerial photographs.

2.5. Likelihood

The final ingredient in the Bayesian model of Eq. (2) is the likelihood $p(D|\theta, I)$, which describes the probability density of observing the data D given some camera parameters θ under the model and prior information I . In our case, the data D is the sequence of 2D projected coordinates (u_p, v_p) for each control point p in the image, corresponding to identified (x_p, y_p, z_p) in 3D space.

Given θ , we may form the 2D projection $(\tilde{u}_p, \tilde{v}_p)$ in the image plane of a 3D control point (x_p, y_p, z_p) using the camera model:

$$(\tilde{u}_p, \tilde{v}_p) := \mathbf{C}(x_p, y_p, z_p; \theta). \quad (15)$$

We can then use the distance between the reprojected control point $(\tilde{u}_p, \tilde{v}_p)$ and the 2D control point (u_p, v_p) as a source of information about how likely it is to observe (u_p, v_p) if the actual underlying point location were $(\tilde{u}_p, \tilde{v}_p)$ as prescribed by the projection model and θ .

It is useful to think of control points as well-defined (if arbitrary) points in the real world which could conceptually be pinpointed with an on-site survey to, say, centimetric accuracy. This could be for example the corner angle of a building at ground level, the exact center point of circular roundabout, or the center of the base of a summit cross. In this picture, there are then two sources of statistical uncertainty impacting the 2D/3D matches in the data D .

Firstly, the user-provided placement of the 2D control points (u_p, v_p) in the image is approximate: point picking generally involves some level of uncertainty, for instance when determining the exact location of a summit point along an elongated ridge. The accuracy of 2D point picking depends on many factors, including image resolution, blurriness, perspective, control point visibility, etc. Given an image, the user is generally able to infer a best-guess location for a well-defined control point, together with some associated point-dependent uncertainty. We therefore associate to each 2D control point p in the image plane a user-specified uncertainty radius e_p in pixels, which represents the user's assessment of the uncertainty in the point picking process. We choose to use 90% credible regions, so that when placing a 2D control point and its associated confidence radius, the user should expect the actual feature to lie within the designated area with 90% confidence. While there is an amount of subjective probability estimation involved in this procedure, we find that it works well in practice, and that this threshold of 90% is rather intuitive to work with.

Secondly, the 3D locations (x_p, y_p, z_p) of well-defined control points

can only be determined up to some accuracy. Some control points could potentially be surveyed on-site to centimetric accuracy with differential GPS, whereas coordinates derived from maps typically only provide accuracies of a few meters at best. Therefore, we also account for a 3D uncertainty radius E_p in meters for 3D points coordinates. In the case studies presented in this work, all 3D points come from sources of similar accuracy, and we typically set $E_p = 2 \text{ m}$ independent of p as the 90% credible region for control points.

In practice, for a given 2D/3D control point pair, we combine the 2D and 3D uncertainties e_p and E_p by adding them in quadrature into a single total 2D pixel uncertainty. A 3D object of size E_p located at control point p will project onto the image plane as a feature of approximate size $f_{\text{px}} E_p / d_p$ pixels, where d_p is the longitudinal distance between the object and the camera, and f_{px} is the camera focal length expressed in pixels. As a result, we compute the total uncertainty radius ϵ_p in pixels for control point p as

$$\epsilon_p := \sqrt{e_p^2 + (f_{\text{px}} E_p / d_p)^2}. \quad (16)$$

Note that f_{px} and d_p are a function of the camera parameters θ , and may therefore be used in the likelihood computation where θ is assumed to be known. Eq. (16) combines both uncertainties into a single pixel uncertainty ϵ_p in the image plane, as it is in image space that the likelihood evaluation takes place. The addition of the variances in (16) can be seen as an approximation of the convolution of the distributions for the point picking and 3D coordinates uncertainties, which arises when marginalizing over possible exact (error-less) 3D control point locations.

We assume that all points are statistically independent, in the sense that all 2D user placement errors and 3D uncertainties are all uncorrelated. Under this assumption, we may write:

$$p\left(D \mid \theta, I\right) = p\left(\left(u_1, v_1\right), \dots, \left(u_p, v_p\right) \mid \theta, I\right) = \prod_{p=1}^n p\left(u_p, v_p \mid \theta, I\right). \quad (17)$$

To evaluate $p(u_p, v_p | \theta, I)$ for control point p , we first apply the full camera model \mathbf{C} of Section 2.3 to compute the projected coordinates $(\tilde{u}_p, \tilde{v}_p) = \mathbf{C}(x_p, y_p, z_p; \theta)$. The likelihood is then computed by comparing the modeled point $(\tilde{u}_p, \tilde{v}_p)$ with the user-specified data point (u_p, v_p) .

Because we model our credible regions as circles, we assume that the likelihood only depends on the error distance

$$r_p := \sqrt{(u_p - \tilde{u}_p)^2 + (v_p - \tilde{v}_p)^2} \quad (18)$$

between the data point and the modeled point.

Given a user-provided point (u_p, v_p) and associated total 90% confidence radius ϵ_p obtained from Eq. (16), a simple likelihood model would be to take (u_p, v_p) normally distributed around $(\tilde{u}_p, \tilde{v}_p)$, with covariance $\sigma_p^2 \mathbf{I}_2$, where $\sigma_p \approx 0.466 \epsilon_p$ is set so that $p(r_p < \epsilon_p) = 0.9$ (i.e. 0.466 is the inverse of the Rayleigh or χ_2 quantile for 90%). However, we find that using a Gaussian likelihood can be problematic in some cases, because it is not robust to outliers (Sheynin, 1995; Fonseca et al., 2008). Multiple methods exist to make Bayesian methods robust (see e.g. Sivia and Skilling, 2006). Approaches to robust camera calibration based on Bayesian methods have been presented, for example by (Krahnstoever and Mendonca, 2005) using mixtures of Gaussian and uniform likelihoods to account for potential outlier points.

In this work, we follow the simple and well-established practice of replacing the normal distribution with a Student- t distribution, which has similar overall properties but features heavier tails, and is therefore more susceptible of accommodating outlier points (Shah et al., 2014). We therefore define, up to a normalization constant independent of θ :

$$p\left(u_p, v_p \mid \theta, I\right) \propto \eta_p^{-2} \left(1 + \frac{r_p^2}{\eta_p^2} \nu^{-1}\right)^{\frac{\nu+1}{2}}, \quad (19)$$

where r_p is the error radius of (18) for point p , $\nu > 0$ is a fixed

parameter, and η_p is a scaling radius which will depend on θ . The parameter ν controls the heaviness of the tails, with $\nu \rightarrow +\infty$ approaching the normal distribution, and lower values of ν producing heavier tails. In practice, values of $3 \lesssim \nu \lesssim 10$ are often recommended for robust inference (see e.g. Gelman et al., 2013, Chap. 17); we adopt a fixed value $\nu = 5$. For this value of ν , we find for the scaling factor $\eta_p \approx 0.304 \epsilon_p$ by requiring that $p(r_p < \epsilon_p) = 0.9$. Note that the overall proportionality constant in (19) depends on ν and η_p only and not on any parameter of θ . In practice, this normalization is not important for the sampling process described in Section 4.1.

3. Inverse perspective and uncertainties

We now discuss the inverse perspective step, in the form of the back projection, which is needed to reconstruct the area of a landscape feature from its outline in the image. In Section 3.1, we first describe the back projection process through which we obtain the area S from the camera calibration results. To account for uncertainties in both the polygon tracing process as well as DEM elevations, we derive in 3.2 the posterior distribution of S , assuming imperfect knowledge of the polygon and digital elevation model. Finally, in Sections 3.3 and 3.4, we describe our statistical models for the polygon and DEM uncertainties respectively.

3.1. Polygon back projection

Polygon back projection consists in obtaining the 3D world coordinates of each vertex of a polygon P drawn on the image plane, given *fixed and known* camera parameters θ , polygon P , and DEM H . Each vertex of P is back projected independently in a two-step process.

First, we determine the world coordinates of the corresponding ray, i.e., the unit vector pointing from the camera center to the image pixel of the considered polygon vertex. To this end, we first undo the lens distortion to recover the vertex pixel coordinates for a perfect lens. This requires inverting (6) and (7) jointly, which we do numerically using a small number of Newton–Raphson iterations. The resulting pixel is then interpreted as a ray, cast from the pinhole camera’s center position through the corresponding pixel position in the 2D image plane. The ray orientation in 3D world coordinates is reconstructed from the camera angles.

In a second step, we search along the ray direction for the first intersection point with the ground, whose local elevation is given by the DEM H . The intersection point is then the sought back projected 3D vertex. By construction, this 3D vertex is the unique 3D point closest to the camera and located at the surface of the DEM which will project exactly onto the corresponding polygon 2D vertex in the image plane.

3.2. Feature area posterior distribution

We now come back to the problem of determining the posterior distribution of the surface area S of a polygon traced on the image. In this work, we account for uncertainties on the 2D/3D matches D (through the Bayesian camera calibration process), on the tracing of the polygon P , and on the digital elevation model H . We stress that the following derivation can easily be extended to other sources of uncertainties, provided that relevant statistical models are provided.

The uncertainties on the control point matches is described by the posterior PDF of the camera parameters from the camera orientation and calibration problem, $p(\theta|D, I)$. To account for the polygon and DEM uncertainties, let us note \tilde{P} the polygon traced by the user, and \tilde{H} the DEM used for the study. The tilde notation highlights the fact that both \tilde{P} and \tilde{H} are imperfect representations of the true, exact underlying polygon P and DEM H , which are unknown and inaccessible. The difference between P and \tilde{P} is caused by errors in the manual user tracing process, whereas \tilde{H} differs from H because of deviations between the DEM and actual “true” elevation map.

Given the control point matches D , the user-traced polygon \tilde{P} , the DEM \tilde{H} , and prior information I , the posterior of S may be informally expressed as

$$p(S|D, \tilde{P}, \tilde{H}, I) = \int d\theta \int \mathcal{D}P \int \mathcal{D}H p(S, \theta, P, H | D, \tilde{P}, \tilde{H}, I) \quad (20)$$

where we have marginalized over all camera parameters θ , as well as all the possible *exact* polygons P and DEMs H compatible with \tilde{P} and \tilde{H} . The integrals $\int \mathcal{D}P$ and $\int \mathcal{D}H$ run over the spaces of all possible exact polygons P and DEMs H with appropriate probability measures; we can avoid the complication of formally defining these integrals by using Monte-Carlo sampling, whose application to the posterior of S is described in Section 4.2.

Splitting the joint PDF on the right-hand side using conditional probabilities, we get:

$$p(S, \theta, P, H | D, \tilde{P}, \tilde{H}, I) = p(S|\theta, P, H, I) p(\theta, P, H | D, \tilde{P}, \tilde{H}, I) \quad (21)$$

where we have used the fact that in the back projection process, S depends on P and H only. Further splitting the right-hand product term using conditional probabilities on P and H , noting that the posterior on θ directly depends on D , \tilde{H} and I only,⁵ and assuming that the statistical uncertainties on the polygon tracing and the DEM are independent, the integrand of (20) becomes:

$$\frac{p(S|\theta, P, H, I)}{\text{Back-proj.}} \times \frac{p(\theta|D, \tilde{H}, I)}{\text{Calibration}} \times \frac{p(P|\tilde{P}, I)}{\text{Poly. uncert.}} \times \frac{p(H|\tilde{H}, I)}{\text{DEM uncert.}} . \quad (22)$$

This expression collects all the sources of statistical uncertainty accounted for in the model; it is easy to see how it can be expanded in a similar fashion to account for more sources of errors.

Because the back projection process yields a fully determined area S given θ , \tilde{P} , \tilde{H} and I , as described in 3.1, the probability density of S simply writes:

$$p(S|\theta, P, H, I) = \delta(S - S(\theta, P, H, I)), \quad (23)$$

where δ is the Dirac delta function. Note that although (20), (22) and (23) look unwieldy, they lend themselves well to sampling $p(S|D, P, H, I)$; in practice we use Monte-Carlo sampling to probe the posterior of S , which we will describe in 4.2.

3.3. Characterization of polygon tracing uncertainties

Tracing uncertainties arise from the user’s approximate placement of the polygon vertexes on the picture. This results in the user-input polygon \tilde{P} differing from the actual underlying error-free polygon P . The discrepancy is described statistically by specifying in (22) a PDF $p(P|\tilde{P}, I)$ on the space of 2D point coordinates of P . Note that tracing errors are more intuitively specified by the reversed probability density $p(\tilde{P}|P, I)$, that is, the probability density of observing a given traced polygon \tilde{P} suffering from tracing errors, knowing the true underlying polygon P . Formally, the two densities $p(P|\tilde{P}, I)$ and $p(\tilde{P}|P, I)$ are related through Bayes’ theorem; in the following we will assume for simplicity that $p(P|\tilde{P}, I) = p(\tilde{P}|P, I)$, which in our following model holds in the limit of small errors.

Providing an accurate description of $p(P|\tilde{P}, I)$ is a potentially complicated problem, as it requires investigating in detail sources of errors in the tracing process. In this work, we settle for a simple model

⁵ Note that formally, the posterior on θ should be computed as $p(\theta|D, H, I)$, and not $p(\theta|D, \tilde{H}, I)$. However, the DEM only enters this posterior in the case of terrestrial images through the prior for $z_c(x_c, y_c)$, for which we choose a vertical uncertainty (5 m) which is typically comparable to the estimated RMSE between H and \tilde{H} (4 m). Therefore, in this case the simplifying approximation $\tilde{H} \approx H$ is justified, and allows decoupling the camera estimation from the DEM Monte-Carlo sampling.

in which \tilde{P} and P have the same number n of vertexes. We define a vertex normal 2D vector \mathbf{n}_j at each vertex j of the polygon \tilde{P} , and restrict ourselves to errors which displace vertexes along \mathbf{n}_j , as these types of errors will have the most impact on the shape of the polygon and its total surface area. Errors are therefore characterized by the collection of n real numbers

$$\lambda_j := (\mathbf{v}_j - \bar{\mathbf{v}}_j) \cdot \mathbf{n}_j, \quad j = 1, \dots, n, \quad (24)$$

with $\lambda_j = 0$ yielding $\mathbf{v}_j = \bar{\mathbf{v}}_j$. We assume that the error amplitudes λ_j follow a multinormal distribution with zero mean and covariance matrix

$$\text{cov}(\lambda_j, \lambda_k) = \sigma^2 e^{-d_{jk}/\ell}, \quad (25)$$

where d_{jk} is the shortest path distance between vertexes j and k along the polygon perimeter, and ℓ is a correlation length scale that we chose to be 1/20 of the total perimeter of the polygon \tilde{P} . This ensures that the polygon can have about 20 continuous “segments” of vertexes whose errors are roughly independent. The σ parameter describes the overall scaling of the errors, and in the rest of this work we take $\sigma = 1$ pixel, corresponding to a user tracing the true underlying polygon within about ± 1.6 pixels in the normal direction 90% of the time.

3.4. Characterization of DEM uncertainties

Because the back projection procedure consists in intersecting camera rays with the DEM, unknown deviations of the elevation model \tilde{H} from actual terrain elevation H will result in additional uncertainties on the feature area of interest. In this section, we describe a simple model which plays for a DEM H an analogous role to the uncertainty model 3.3 for the polygon P . The goal of this model is to generate possible true underlying DEMs H given a DEM \tilde{H} with errors, i.e. to sample the $p(H|\tilde{H}, I)$ term in (22). Note that here again, we conflate $p(H|\tilde{H}, I)$ and $p(\tilde{H}|H, I)$ into a same DEM uncertainty process. Just like for the polygon, this is justified in the limit of small perturbations. From a practical standpoint, \tilde{H} constitutes our best available elevation information (see Wechsler and Kroll, 2006, for a discussion of this point in a similar context).

The accuracy of digital elevation models has been extensively studied in connection to various problems, e.g. for viewshed and intervisibility problems (Fisher, 1991), automatic feature extraction (Lee, 1992), the creation of DEMs from multiple data sources (Kyriakidis et al., 1999), or the determination of local topographic parameters (Wechsler and Kroll, 2006). Various methods have been proposed to model realistic autocorrelated DEM uncertainties based on accurate elevation measurements at ground control points, often relying on stochastic or random field approaches (Lee, 1992; Hunter and Goodchild, 1997; Fisher, 1998; Kyriakidis et al., 1999; Wechsler and Kroll, 2006).

Deviations between an approximate DEM \tilde{H} and true elevation map H are often first described using simple metrics, which capture the globally averaged properties of the error. A commonly used indicator is the root mean square error (RMSE) (e.g. Li, 1988; Fisher, 1998):

$$\text{RMSE} := \sqrt{\frac{1}{N} \sum_i (\tilde{H}_i - H_i)^2}, \quad (26)$$

for which estimates are usually provided in the DEM product documentation.

Global error metrics are insufficient however to fully characterize DEM errors. The magnitude of errors has been found correlate with a number of local terrain properties (Carlisle, 2005), including in particular slope (Liu and Jezek, 1999; Carlisle, 2005), ruggedness (Kyriakidis et al., 1999; Liu and Jezek, 1999), and possibly local terrain elevation (Kyriakidis et al., 1999). The documentation for our DEM, for instance, indicates an RMSE error varying from 1 m in plains to 4 m in high-altitude rugged alpine areas.

Beyond the magnitude of DEM errors measured as a space-varying RMSE, DEM errors are also spatially autocorrelated (see e.g. Fisher, 1991; Fisher, 1998; Kyriakidis et al., 1999; Lee, 1992; Holmes et al., 2000; Carlisle, 2005). Based on a semivariogram analysis, some authors find DEM error autocorrelations to be well described by a two-component model: a short-scale correlation, over separations of 20 m to 125 m, and larger-scale correlations at kilometric scales (Holmes et al., 2000; Haneberg, 2006). It has also been suggested that the error autocorrelation scale is itself position-dependent, based on the local topography, with the autocorrelation length varying with the scale of local terrain features (Liu and Jezek, 1999).

In this paper, we opt for a simple model which still attempts to capture all these elements. For simplicity, we assume that our DEM \tilde{H} is unbiased, in the sense that the average of each H_i over all statistical realizations of the error is equal to the true elevation H_i at each DEM node i , however it is very easy to relax this assumption if prior information on the DEM supports doing so. We adopt an ad hoc model with locally-varying RMSE $\sigma(x, y)$, as well as locally-varying correlation scale length $\lambda(x, y)$. The maps for σ and λ are constructed from the DEM \tilde{H} which describes the elevation map $z(x, y)$. Since our DEM specifications document different values of the RMSE for low-altitude and flat areas on the one hand, and high-altitude and alpine terrain on the other, we propose to define a transition map $\xi(x, y) \in [0, 1]$, with $\xi \approx 0$ in plains, and $\xi \approx 1$ in mountainous areas. We construct this transition map from the local elevation $z(x, y)$, and a proxy quantity $q(x, y)$ for local terrain ruggedness.

We start by computing a map of local terrain ruggedness q_i at each DEM node i , roughly following the definition⁶ of (Kyriakidis et al., 1999):

$$q_i := h^{-1} \sqrt{\frac{1}{8} \sum_{j \rightarrow i} \left(z_j - \frac{1}{9} \sum_{k \rightarrow i} z_k \right)^2}, \quad (27)$$

where the sums on j, k run over the 3×3 square of nodes immediately around node i , and h is the horizontal DEM resolution (grid cell size) in meters. Therefore, q_i is proportional to the local standard deviation of DEM elevations in the 3×3 square around i . Note that $q \geq 0$ is dimensionless, and is closely related to the local terrain slope. Regions of terrain with low q will be found in flat, featureless plains and plateaus, whereas steep slopes or very irregular terrain will result in large ruggedness values.

From the ruggedness map q , we construct an ad hoc transition map ξ as:

$$\xi = \tanh \left[Aq + B \min \left(\max \left(\frac{z - 2000 \text{ m}}{3000 \text{ m} - 1000 \text{ m}}, 0 \right), 1 \right) \right]. \quad (28)$$

The term in z behaves as a smooth linear ramp between 0 at $z = 2000$ m and 1 at $z = 3000$ m, saturating a 0 below and 1 above respectively. A and B are positive weights which control the relative scaling of the ruggedness and elevation in the transition map; in all of the following we set $A = B = 1$. The tanh function ensures that the resulting ξ is always in the interval $[0, 1]$, introducing a nonlinear saturation for large values of its arguments, but responding linearly to small arguments. Fig. 3 illustrates the transition function ξ on the Mont Blanc study area, with the source DEM z shown in panel (a), and the resulting transition function ξ in panel (b). The altitude cutoffs were chosen to broadly discriminate between sub-alpine and alpine elevations; the resulting ξ map is smoothly varying across the whole area of study, and sensitive to both altitude and local terrain ruggedness.

We then construct maps for the RMSE σ and error autocorrelation

⁶ The factor of $8 = 9 - 1$ appearing in our formula, corresponds to the scaling factor for an unbiased estimation of the standard deviation. Compared to (Kyriakidis et al., 1999), the extra $1/h$ normalization factor ensures that our definition of q is independent of the DEM horizontal resolution h .

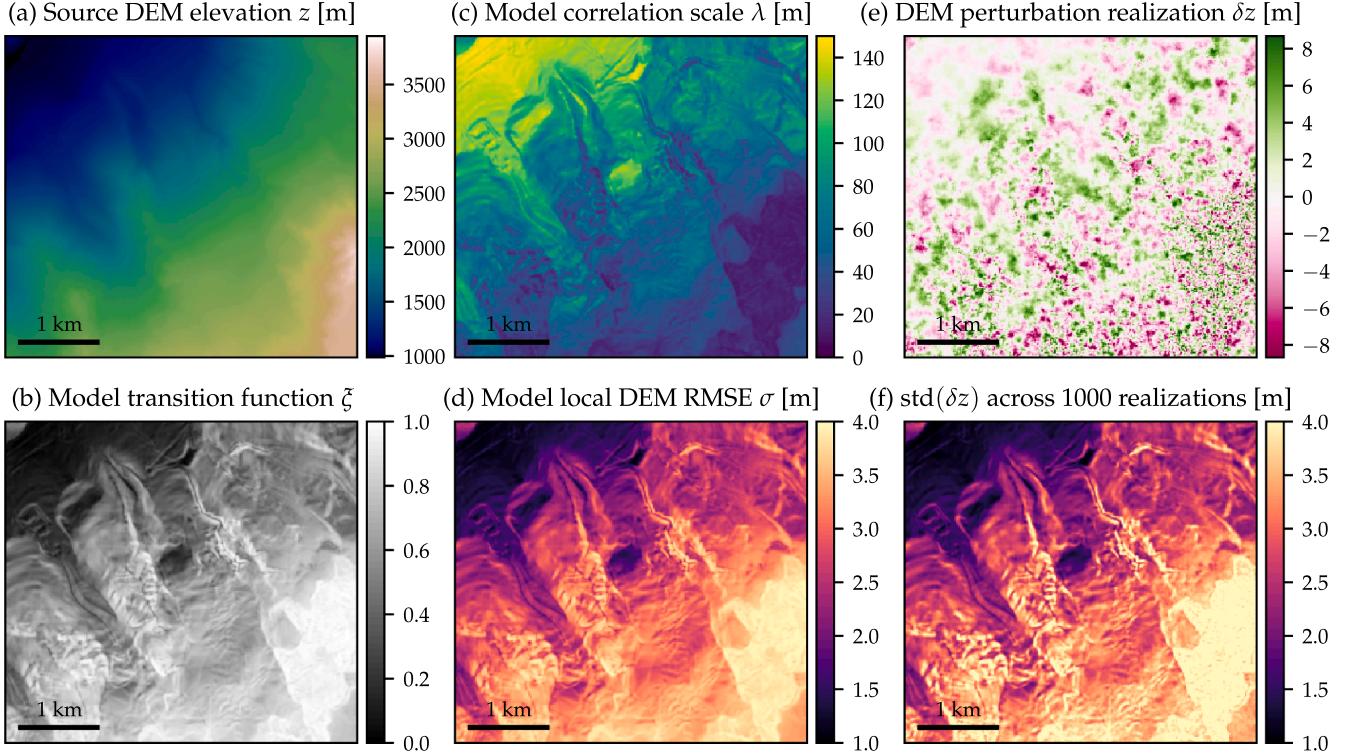


Fig. 3. Illustration of the DEM perturbation generation process. Panel (a) shows the DEM elevation map z around the area of interest for the Mont Blanc study. From the elevations z , the transition function $\xi \in [0, 1]$ shown in panel (b) is constructed so that $\xi \approx 0$ in regions of smooth and low altitude terrain, and $\xi \approx 1$ in rugged and mountainous areas. Models for the error correlation scale λ and RMSE σ (panels (c) and (d) respectively) are then derived from ξ based on DEM documentation and other studies of DEM errors. From the λ and σ maps, a random realization δz of an elevation perturbation is then generated as described in the text. The resulting field is presented in panel (e), clearly showing the spatially varying correlation scale. Panel (f) shows the standard deviation of δz across 1000 random realizations, illustrating that the generated DEM perturbations closely follow the prescribed RMSE map σ .

scale λ based on the transition function. σ is constructed by linearly interpolating between the documented DEM RMSE errors in flat plains and high alpine areas according to the transition map:

$$\sigma := (1 - \xi) \times 1 \text{ m} + \xi \times 4 \text{ m}. \quad (29)$$

Panel (d) of Fig. 3 illustrates the resulting map for σ as specified in the model.

For the DEM perturbation autocorrelation range λ , we adopt ad hoc values between 20 m and 150 m, based on the small-scale component observed by (Holmes et al., 2000; Haneberg, 2006) in their semivariogram analysis. We note that the other, multi-kilometer-scale correlation component found by these studies is unlikely to significantly impact our study because the terrain features we study are typically at most kilometer-sized.

Our prescription for the spatial dependence of λ follows (Liu and Jezek, 1999), who find that the autocorrelation length of the DEM errors varies according to the scale of local terrain features. In rugged areas with small-scale terrain variation, we set $\lambda = 20$ m. In smooth areas, terrain varies only on larger scales, and we correspondingly adopt a larger value of $\lambda = 150$ m. The map for λ is therefore constructed simply using the transition map as:

$$\lambda := (1 - \xi) \times 150 \text{ m} + \xi \times 20 \text{ m}. \quad (30)$$

Panel (c) of Fig. 3 illustrates the resulting map for λ corresponding to this model.

Finally, we construct random realizations of DEM perturbations δz with local standard deviation and local autocorrelation scale length given by the maps σ and λ . This process is based on a random field approach, inspired by the autoregressive technique of Hunter and Goodchild (1997) in which spatially correlated DEM disturbance fields

u are generated by iterating⁷:

$$u \leftarrow \frac{\rho}{4} \mathbf{W} u + N \quad (31)$$

for $0 \leq \rho \leq 1$. \mathbf{W} is the matrix with ones on “rook” neighbor positions, (i.e. the very next horizontal and vertical neighboring DEM nodes) and 0 everywhere else. N is a vector of point-wise independent normal noise realizations, and is fixed across the iteration process. By iterating this process to convergence, Hunter and Goodchild (1997) obtain a random field u with spatial correlations.

In this work, we slightly modify and extend this prescription, in order to generate DEM perturbations $\delta z(x, y)$ with given maps of autocorrelation length $\lambda(x, y)$ and RMSE $\sigma(x, y)$ (i.e. variance σ^2). For a DEM with grid mesh point spacing (resolution) h , we solve the discrete equation⁸ in \tilde{u} :

$$(\mathbf{W} - 4\mathbf{I})h^{-2}\tilde{u} - \lambda^{-2}\tilde{u} = -2\sqrt{\pi}h^{-1}\lambda^{-1}N, \quad (32)$$

⁷ Note that our definition of ρ differs from (Hunter and Goodchild, 1997) by a factor of 4, but otherwise the process is identical.

⁸ This equation can be obtained from (31) by letting $\lambda^{-2} := 4h^{-2}(\rho^{-1} - 1)$ and considering the fixed point of the iteration process in the limit of infinitely fine mesh spacing $h \rightarrow 0$, which yields $\Delta u - \lambda^{-2}u = -4\rho^{-1}h^{-2}N$. This is a linear stochastic partial differential equation (PDE) for a random field u with correlation length λ . The local variance of its (continuous) solution u can be computed analytically to be $\text{var } u = 4\pi^{-1}\rho^{-2}h^{-2}\lambda^2$. Letting $\tilde{u} := \sqrt{\pi}\rho h \lambda^{-1}u/2$ and discretizing this last PDE on the DEM grid yields Eq. (32) with $\text{var } \tilde{u} = 1$, and Eq. (33) therefore produces δz with the desired properties. Note that while this derivation is only formally correct for λ and σ uniform in space, we find that using space-varying $\lambda(x, y)$ and $\sigma(x, y)$ works very well in practice and generates fields with appropriate statistical properties, as illustrated in Fig. 3.

and obtain the DEM perturbation with

$$\delta z := \tilde{u} \sigma. \quad (33)$$

The resulting $\delta z(x, y)$ corresponds to a random DEM perturbation with approximate correlation scale length $\lambda(x, y)$ and RMSE $\sigma(x, y)$. Each random realization of δz is obtained by solving Eq. (32) with a different independent realization of the white noise right-hand-side N . In practice, we solve (32) using an efficient algebraic multigrid solver (Olson and Schroder, 2018). Panel (e) of Fig. 3 shows a single realization of a random δz , on which both the point-dependent variance and spatially-varying correlation scale is visible. Finally, panel (f) presents the pixel-wise standard deviation of δz across 1000 random realizations; the resulting standard deviation map follows the map of the RMSE σ of panel (d) closely, illustrating that the random field approach successfully reproduces the local variance prescribed by the model.

To conclude, coming back to the problem of characterizing DEM uncertainties, recall that we want to produce realizations H that sample the $p(H|\tilde{H}, I)$ term in (22). For each δz realization obtained from \tilde{H} by the process described in this section, we obtain a sample from $p(H|\tilde{H}, I)$ by constructing the DEM $H = \tilde{H} + \delta z$.

4. Implementation and MCMC sampling

The first step towards implementation is evaluating the Bayesian posterior on camera parameters θ . We have by now specified the full prior (Eq. (11) with terms discussed in Section 2.4), and likelihood using the camera model and Eqs. (17) and (19). We can therefore evaluate the posterior probability density $p(\theta|D, I)$ using (2) for any value of the parameter θ .

4.1. MCMC sampling of the posterior on camera parameters

Because the parameter space of θ is high-dimensional, it is intractable to explore the posterior PDF (2) by gridding parameter space. To overcome this common problem in Bayesian inference, Monte Carlo Markov Chain (MCMC) methods have been developed. Given a way to evaluate a probability distribution function over a possibly high-dimensional space, MCMC methods produce individual samples from a random walk whose stationary distribution asymptotically converges to the chosen PDF.

In this work, we use the `emcee` implementation (Foreman-Mackey et al., 2013) of the affine-invariant MCMC ensemble sampler proposed by Goodman and Weare (2010) to sample the posterior PDF for the camera parameters. This method relies on the mixing of a number N_c of Markov chains (called *walkers*), which are advanced together in a step by step process.

At the beginning of the MCMC sampling process, we initialize the N_c walkers at values of θ randomly drawn from the prior for θ ; this ensures that the walkers are scattered across the plausible parameter space. The random walk typically takes some number of steps τ to forget about its own history, characterized by the chain's autocorrelation time. We therefore first run the process for a number N_w of warmup steps from the initial positions, and the corresponding samples are discarded. After warmup, the chains are advanced by N_s additional steps for sampling. At the end of the sampling process, we obtain a total of $N_c N_s$ sampled values of θ , with N_s values for each of the N_c chains.

Markov chains should be checked for stationarity and convergence before inferring properties of the posterior from the samples. We first visually check for convergence, by plotting the sequence of MCMC chain samples. More quantitatively, we also estimate the autocorrelation time τ of the Markov chains; in our tests, typical values of τ at stationarity are around 130–160. We typically use $N_c = 400$ walkers, each sampling $N_s = 20000$ samples, after $N_w = N_s/3$ warmup steps. This results in about $N_s/\tau \sim 140$ statistically independent samples per chain. Since we use hundreds of chains over many autocorrelation times, we achieve good sampling of the posterior distribution.

Whenever a point estimate is required for θ , different prescriptions can be computed from the MCMC samples (see Hogg et al., 2010). In the rest of this work, point estimates are computed as the median of all sample values from the MCMC sampling. Note that in cases where the posterior distribution for θ is strongly skewed or multi-modal over the multidimensional θ parameter space, this choice could potentially result in a value of θ that is very unlikely (i.e. that lies in an area of very low posterior probability density); it may then be more appropriate to compute a single point estimate of θ by maximizing the posterior probability density using traditional nonlinear optimization methods.

4.2. Sampling of the posterior of S

Once we have obtained samples θ_i characterizing the posterior $p(\theta|D, I)$, we can derive the posterior on the surface area $p(S|D, \tilde{P}, \tilde{H}, I)$ based on (20), (22), (24) and (25), and generate posterior samples for the area S by the following process:

1. Samples θ_i are first generated using MCMC, sampling the posterior PDF of the camera parameters from the camera calibration process,
2. For each sample θ_i :
 - (a) an independent polygon P_i is drawn randomly, by drawing all the λ_j using (25) and then constructing P_i from \tilde{P} using (24),
 - (b) a DEM elevation map H_i is constructed from \tilde{H} and a random realization δz_i of a DEM perturbation as described in 3.4,
 - (c) The resulting $S_i = S(\theta_i, P_i, H_i, I)$ is finally computed using the deterministic back projection process described in 3.1.
3. The S_i obtained by this process then sample the PDF of S according to (20)–(22).

Note that since 2b is a costly operation requiring solving an elliptic problem on the DEM grid, we typically only generate a new random H_i only once every 10 samples θ_i , and reuse values of H_i in-between. This significantly reduces the overhead of sampling DEM perturbations, and does not modify the sampling distribution of the H_i .

5. Validation

In this section, we setup validation case studies to demonstrate the technique presented in this paper. We use a combination of photographs, both aerial and terrestrial oblique of different origins and quality, to assess key aspects of our method.

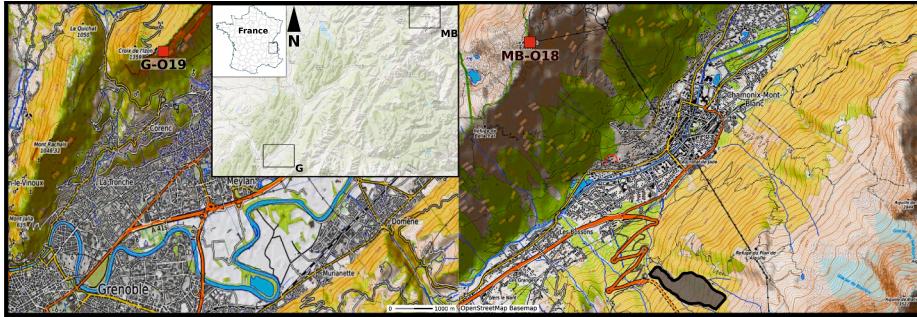
First, we consider the problem of camera orientation with calibration, by fitting images for which the camera location is known approximately, and comparing the obtained posterior on the camera position to the available photograph information.

Based on these fits, we discuss goodness-of-fit metrics for the model, such as control point residuals, and investigate an intrinsically Bayesian test of the model in the form of posterior predictive checking.

Finally, we consider the inverse perspective problem of reconstructing the surface area of geological features from single images. For a specifically geological feature in the Mont Blanc massif, we use back projection to determine the area of the zone from different images and compare the accuracy and systematic effects present in the result. We show that some ingredients of the model, such as accounting for lens distortions, can have a significant impact on the accurate determination of features by back projection when applied to diverse non-precision photographs.

5.1. Presentation of validation case studies and data

We start by describing the input data used in this study, focusing on the images, digital elevation model, and construction of 2D/3D control point matches.



5.1.1. Case studies overview

Fig. 4 presents an overview map of the two different case study areas.

Grenoble area ("G"). This case study is located near the Grésivaudan valley, a broad, flat expanse flanked West and East by the massifs of Chartreuse and Belledonne. As a result, terrestrial oblique pictures feature a mixture of coplanar control points in the valley and remote background points on the background mountains, making it a good test for camera orientation. The area is shown in Fig. 6 showing image G-O19.

We do not perform any back projection analysis on this study, instead focusing on camera orientation and calibration. We use altitude range priors for all photographs in this study, and therefore do not require any digital elevation model of the area.

Mont Blanc massif ("MB"). This case study is located in the Mont Blanc massif, near Chamonix-Mont-Blanc, France. The studied area is located on steep, rugged terrain, that is however considered stable according to the French Land and Forestry Service (RTM ONF, 2019). In particular, the studied zone is not indicated as being prone to landslides (RTM ONF, 2002), and its terrain is therefore likely relatively stable across the years, making it suitable for tracing geological features.

For this case study, we first perform camera calibration and orientation, and also test the inverse perspective problem of calculating the planimetric area of a polygon drawn on the input images, which traces a geological feature of a geometry and configuration relevant for our glacier studies. The overall area is visible in Fig. 5 showing image MB-O18.

5.1.2. Photographs and photographic information

For both case studies, we use publicly available photographs from aerial surveys by the French Institut Géographique National (IGN), as well as terrestrial oblique pictures taken by the authors specifically for the study with different low-grade cameras. Photographs will hereafter be labeled by the prefix of their case study (G or MB), followed by either A## or O##, for "Aerial" and "terrestrial Oblique" photographs,



Fig. 5. Terrestrial oblique image MB-O18 of the Chamonix valley, showing the general terrain configuration, the control points for camera orientation and calibration, as well as the polygonal feature used for the inverse perspective validation.

Fig. 4. Overview map for the two case studies. Geographical coordinates of the center of G and MB zones are N45°12'17"E5°47'18" and N45°55'1"E6°52'17" respectively. Red squares labeled MB-O18 and G-O19 show the camera location for the terrestrial oblique pictures of each study. For the Mont Blanc study, the gray polygon delineates the landscape feature used in the back projection validation section. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

respectively, and where ## corresponds to last two digits of the year the photograph was taken. All photographs used and some of their information are collected in Table 1.

Aerial photographs. All aerial photographs were professionally acquired and digitized by IGN over the course of airborne survey missions, and are publicly available (I.G. National, 2019). The images we studied come from two different analog imaging systems: prior to and including the MB-A93 image, a Wild Heerbrugg RC10 camera system equipped with an Universal Avigoni lens was used, all later images were acquired using a Zeiss RMK TOP 15 aerial survey camera system equipped with a Zeiss Pleonon A3 lens. The plates were scanned by IGN, at resolutions between 20 and 30 μm per pixel (see Table 1).

All aerial images feature fiducial marks from the original photographic plates. We therefore use the tighter aerial priors described in Section 2.4 for lens distortion, lens principal point and pixel aspect ratio.

For some of the missions, we have obtained lens calibration reports from IGN, which provide us with the exact calibrated focal length of the lens, as well as its exact distortion profile. However, for our validation setup, we choose not to set the focal length prior too tightly in order to test the capability of our method to reconstruct the parameters for unknown cameras, as is typically the case for archival or non-scientific images. In addition, this tests the Bayesian method's ability to deal with the additional variance, and with the degeneracy between the depth direction and the focal length.

For a subset of the aerial photographs, some information on the exposure station is available from the image's photographic data panel. These images are marked with “*” in Table 1, and are used to validate the camera orientation solution. Unfortunately, this information display is not very accurate: the exposure station is only displayed to 0.01 arcminute in latitude and longitude. While the RMK TOP computer system records accurate flight path information in order to reconstruct the exact exposure locations once on the ground (Hobbie, 2010), we have not been able to obtain better exposure station information on these images at this time. The aircraft indicated altitude is also shown on the image plate, however, it is difficult to interpret as a height without more information due to its dependence on barometric altimeter calibration and precise atmospheric conditions.

Terrestrial oblique photographs. The two terrestrial oblique photographs, one in each study, were taken by the authors, using consumer smartphone cameras, to test the method with lower-grade lenses in conditions similar to pictures taken without scientific intent. For each image acquisition, the camera location was determined on-site using a consumer Garmin eTrex 30 hand-held GPS unit, operating with EGNOS (European Geostationary Navigation Overlay Service) augmentation. Between the GPS unit's own accuracy indication, and published accuracy for EGNOS GNSS (European GNSS Agency, 2017), we assume the location to be accurate to about 3 m accuracy. The camera altitude was determined using a calibrated barometric altimeter.

The two images, MB-O18 and G-O19, are shown in Figs. 5 and 6 respectively.

Table 1

Table of photographs used for the two validation case studies G and MB. For images marked with “ $*$ ”, some exposure station or camera location information is available, and is used for camera orientation validation.

Image	Type	Source and reference	Camera	Focal length	Scan or image resol.	Number of CPs
G-A96 $*$	Aerial	IGN C96SAA0411-1996-F3234-0090	Zeiss RMK TOP 15	153 mm	21 $\mu\text{m}/\text{px}$	22
G-A98 $*$	Aerial	IGN CA98S01232-1998-FD38-1088	Zeiss RMK TOP 15	153 mm	28 $\mu\text{m}/\text{px}$	16
G-A03 $*$	Aerial	IGN CP03000082-2003-FD0038-250-C-2279	Zeiss RMK TOP 15	153 mm	28 $\mu\text{m}/\text{px}$	16
G-O19 $*$	Terr. oblique	Own	Huawei COR-L29 Smartphone	3.8 mm	4608 \times 3456 px	25
MB-A88	Aerial	IGN C3529-0011-1988FD17-0226	Wild RC10	152 mm	21 $\mu\text{m}/\text{px}$	16
MB-A93	Aerial	IGN C93SAA088-1993-FD74-0492	Wild RC10	152 mm	28 $\mu\text{m}/\text{px}$	15
MB-A01 $*$	Aerial	IGN CA02S00032-2001-FD0073-250-C-3163	Zeiss RMK TOP 15	153 mm	28 $\mu\text{m}/\text{px}$	15
MB-A04	Aerial	IGN CP04000702-2004-FD74-C-20000-0817	Zeiss RMK TOP 15	153 mm	25 $\mu\text{m}/\text{px}$	12
MB-O18 $*$	Terr. oblique	Own	ASUS X00HD Smartphone	3.4 mm	4160 \times 3120 px	18



Fig. 6. Terrestrial oblique image G-O19 of the Grésivaudan valley, showing the general terrain configuration and the control points for camera orientation and calibration.

5.1.3. Digital elevation model (DEM)

For the Mont Blanc case study, the digital elevation model used is provided by a regional French agency (Régie de Gestion des Données des Pays de Savoie). It was constructed from 1998 aerial photographs, revised in 2004 through photogrammetric ground control points, and further updated in 2008 and 2010. The DEM is provided to us as rectilinear gridded data with 4 m node spacing (horizontal resolution) in both easting and northing, in the Lambert-93 projection. The documentation reports an estimated vertical root mean square error (RMSE) varying between 1 m in plains and 4 m in rugged and alpine high altitude areas. The DEM for the MB case study is used both in the camera position prior for the MB-O18 image, and in the inverse perspective back projection step in all MB images.

For the associated DEM uncertainties, we use the model described in 3.4, with the exact parameter values mentioned in that section.

For the Grenoble case study, we do not rely on any DEM, since we do not test back projection in this area, and we use altitude range priors on the camera parameter z_c .

5.1.4. Control points and definition of 2D/3D matches

For each image in both case studies, we manually determine 2D/3D control point matches. We proceed by identifying well-defined features on the photograph, such as road crossings, mountain peaks, corners of noticeable buildings, and match them to their known position in 3D world coordinates. This 3D world position may be determined by correlating diverse sources, such as georeferenced aerial imagery, detailed topographical maps, planimetric and altimetric databases such as OpenStreetMap (OpenStreetMap contributors, 2019). Because of uncertainties associated with many of these data sources, we account for a 3D uncertainty $E_p = 2$ m in the likelihood as described in 2.5.

We define the 2D/3D matches by associating to each 3D control point its 2D pixel coordinates in the image. To account for uncertainties in the placement of each 2D point, the user specifies for each of them a 90% confidence radius (see Section 2.5). Examples of 2D points with

their confidence radii are shown in Fig. 7.

5.2. Camera orientation validation

We now present results from the camera orientation and calibration step for our input images for which camera location information is available. We choose to focus on the reconstruction of the planimetric camera location parameters (x_c, y_c) , since they are crucial parameters for which the reference value for our test images is known with the greatest accuracy. The camera altitude z_c is typically less constrained on aerials because of the degeneracy with the focal length, and uncertainties on altitude measurements.

5.2.1. Terrestrial images

Fig. 8 compares the locations of the camera for our two terrestrial images as determined by the method to the values measured on site. The ellipses show 7 isocontours of the MCMC posterior PDF for (x_c, y_c) as described in the figure caption.

We find that the reconstructed and reference position distributions are in good agreement. Qualitatively in Fig. 8, the reconstruction error (offset between the method's median point estimate and the measured (x_c, y_c, z_c) position) is comparable in magnitude to the combined statistical uncertainties on the reconstructed and reference positions, materialized by the PDF contours and the 3 m radius circle respectively.

More quantitatively, the errors on (x_c, y_c, z_c) are (2.0 m, 2.9 m, 10.5 m) and (-5.8 m, 4.0 m, -7.8 m) for the MB-O18 and G-O19 images respectively. Assuming 3 m errors on the measured reference barometric altitudes, the 90% credible interval for the observed errors on z_c in images MB-O18 and G-O19 are 9.7 m and 6.2 m respectively. Therefore, the position reconstruction errors are comparable to the combined statistical uncertainties on (x_c, y_c, z_c) , and intrinsic measurement errors on the reference position.

5.2.2. Aerial images

For aerial images for which exposure station information is available, we can also compare the reconstructed camera location at the instant of exposure to the information provided.

This comparison is summarized in Fig. 9, where we show both the reference position region and the contours of the posterior of (x_c, y_c) . Overall, the reconstructed camera location is within 50 m of the indicated reference position along each direction. However, the interpretation of this test is made more complicated by the fact that we could not have access to precise exposure station information at the exact instant of exposure.

First, as discussed in 5.1.2, the exposure station information is only available to us rounded to 0.01 arcminute in latitude and longitude. As a result, we can only determine the corresponding GPS position within 15–20 m in each direction. In Fig. 9, we attempt to account for this uncertainty by representing the locations compatible with the indicated reference position as a quadrangle, assuming that the data panel information is rounded down.



Fig. 7. Example of definition of 2D control points for the MB (left) and G (right) case studies. The two lower panels are zooms of the upper row. Circles represent the 90% confidence radius, accounting for the user's judgment on point uncertainty.

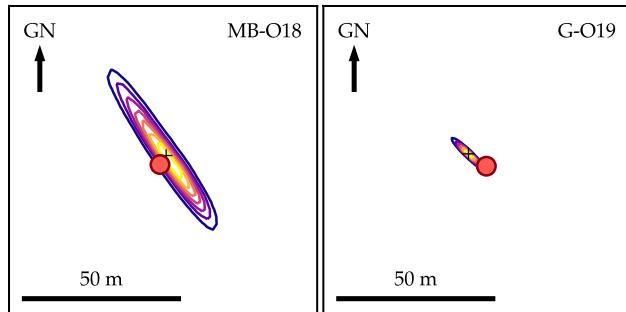


Fig. 8. Camera position validation results for terrestrial oblique photographs. For both images, the ellipses show 7 contours of the MCMC posterior distribution for (x_c, y_c) . The contour levels are $kM/8$, $k \in \{1, \dots, 7\}$, where M is the maximum value of the posterior PDF. The 4th ellipse therefore corresponds to the half maximum level of the PDF. The cross marks the posterior point estimate (median of the MCMC samples), while the circle shows the actual reference camera position measured with a hand-held GPS unit, with an assumed error radius of 3 m. Grid North (GN) of the Lambert-93 projection is up. The posterior and reference point distributions are in very good agreement.

Even at this level of precision however, discrepancies remain in Fig. 9. Over-plotting a trajectory defined by the aircraft's heading (as obtained from the camera orientation solution and our best estimate for the camera mount orientation) shows that the reference GPS position obtained from the plate data panel seems to lag behind the reconstructed aircraft position. The discrepancy distance corresponds to a flight time of approximately 0.2 s to 0.8 s, assuming a typical cruise ground speed of 220 kn as inferred from consecutive exposures. Even though the reconstructed aircraft heading does not always exactly join the reference and reconstructed positions, Fig. 9 shows that the discrepancy is still consistent with the aircraft's possible trajectory assuming moderate crosswinds.

A possible explanation for this delay would be that the data panel display is not refreshed continuously with the exact current GPS location, but rather displays the last position fix updated at the GPS sampling rate, typically set close to 1 Hz for this type of aerial survey (Lapine, 1996). The onboard computerized imaging system would be recording the flight track with precise timing information, and accurate exposure station information can then be computed by interpolation between GPS samples using specialized software on the ground (see Lapine, 1996; Lichten, 2002, as well as Hobbie, 2010 for a historical

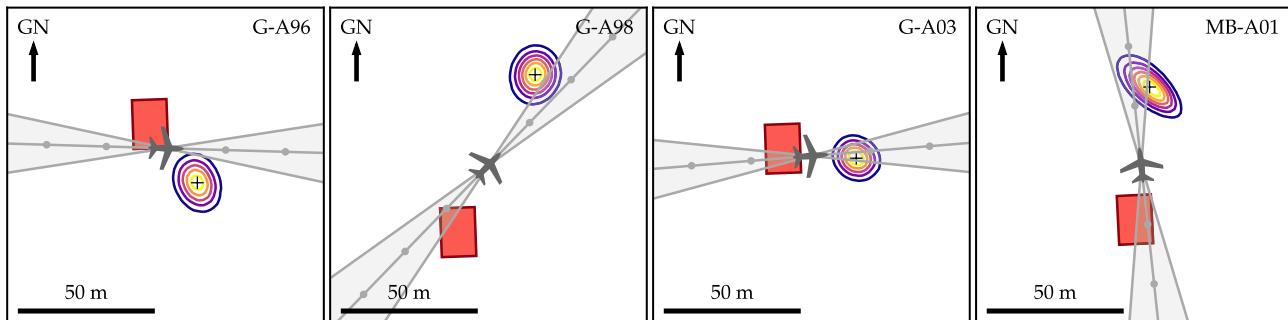


Fig. 9. Camera location reconstruction results for aerial photographs. The contours of the posterior distribution of (x_c, y_c) are shown together with the reference position region (red rectangle) as determined from the data panel display on the photographic plate. Grid North (GN) of the Lambert-93 projection is up. The rectangle spans 0.01 arcminute in latitude and longitude, to reflect the display rounding of the aircraft position. For each image, the probable heading of the aircraft as determined by camera orientation is represented by the median gray line, marked by dots spaced ≈ 0.2 s of flight time apart. The gray lines left and right of the heading line illustrate the corresponding ground track of the aircraft for an indicative crosswind of ± 40 kn. While directly comparing the reference area to the posterior position shows discrepancies of order up to 50 m, these can be explained by a 0.2 s to 0.8 s delay between the GPS position measurement refresh (occurring in the red rectangle), and the actual frame exposure. We discuss this in more detail in the text. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

discussion of the Zeiss RMK TOP system, including of this specific aspect).

While it is difficult to completely settle the question without access to more accurate exposure station information, we find that our method yields camera positions that are broadly consistent with known data, and while discrepancies of a few tens of meters remain on the studied aerial photographs, they can be explained by approximations in the data panel display.

5.3. Goodness-of-fit tests

In this section, we illustrate checks of the goodness-of-fit achieved by the model on the test images from the case studies. We present two tests of the goodness-of-fit between the control points and the model. First, we study the residuals between the reprojected control points and the 2D control points for a single point estimate value of θ ; this tests how well the model is able to fit the control points of an image.

In a second time, we test how well the Bayesian model accounts for variability in the model, and whether the statistical model is consistent with the actual fluctuations observed in the data: this is the purpose of the so-called posterior predictive check.

In all of this section, we work in pixel coordinates in the image plane, and use the notations of Section 2.5.

5.3.1. Reprojection residuals

We now discuss reprojection residuals for a single, fixed image. We first determine a point estimate $\bar{\theta}$ of the camera parameters, which we compute as the median of the MCMC samples θ_i . $\bar{\theta}$ is meant to represent camera configuration typically favored by the posterior distribution from the camera orientation and calibration process.

Given 2D control points (u_p, v_p) and their associated user-provided pixel uncertainties e_p in the image, we compute the reprojections $(\tilde{u}_p(\bar{\theta}), \tilde{v}_p(\bar{\theta}))$ of the corresponding 3D points. We then derive normalized residual vectors with components $((\tilde{u}_p - u_p)/e_p, (\tilde{v}_p - v_p)/e_p)$, which represent the reprojection errors normalized to each control point's uncertainty.

In Fig. 10, we plot these residuals as scaled arrows located at the approximate control point location in the image. The arrow length is proportional to the norm of the error vector (except for very small norms, in which case it is represented as an hexagonal dot). Most error vectors have norm below 1.0, showing that the residual are consistent with the point placement uncertainties e_p . This demonstrates not only that the point estimate $\bar{\theta}$ provides a good fit to the control points in the data D , but also that the errors in the resulting fit are consistent with the user-provided uncertainties on the location of the control points. Aerial images exhibit slightly larger normalized errors compared to oblique photographs, most likely because of their very high resolution, as the uncertainty E_p on the 3D points then becomes more significant than the 2D control point errors e_p .

We also note that the error vectors are isotropic, and do not reveal any strong trend with the location of the control points: this illustrates that there seems to be no systematic tension in the model to reproduce the projected control point locations at $\bar{\theta}$.

5.3.2. Posterior predictive checking

We now present another goodness-of-fit test, which is intrinsically Bayesian and tests the self-consistency of the model's statistical description. This test, called posterior predictive checking, uses the posterior $p(\theta|D, I)$ to describe hypothetical data observations D' under replicated data measurement experiments: in our case, it uses the model to predict control points $D' = \{(u'_p, v'_p)\}$ which we could likely observe as alternates to D , if the control point determination process were to be replicated many times. This prediction relies on the model's inference from the existing control points $D = \{(u_p, v_p)\}$, and in particular takes into account the full Bayesian uncertainty in θ . Posterior predictive checks are commonly used as a self-consistency and goodness-of-fit check in Bayesian modeling (Gelman et al., 1996; Hogg and Foreman-Mackey, 2017; Gelman et al., 2013). The test amounts to sampling from the posterior predictive distribution:

$$p(D'|D, I) = \int d\theta p(D'|\theta, I)p(\theta|D, I), \quad (34)$$

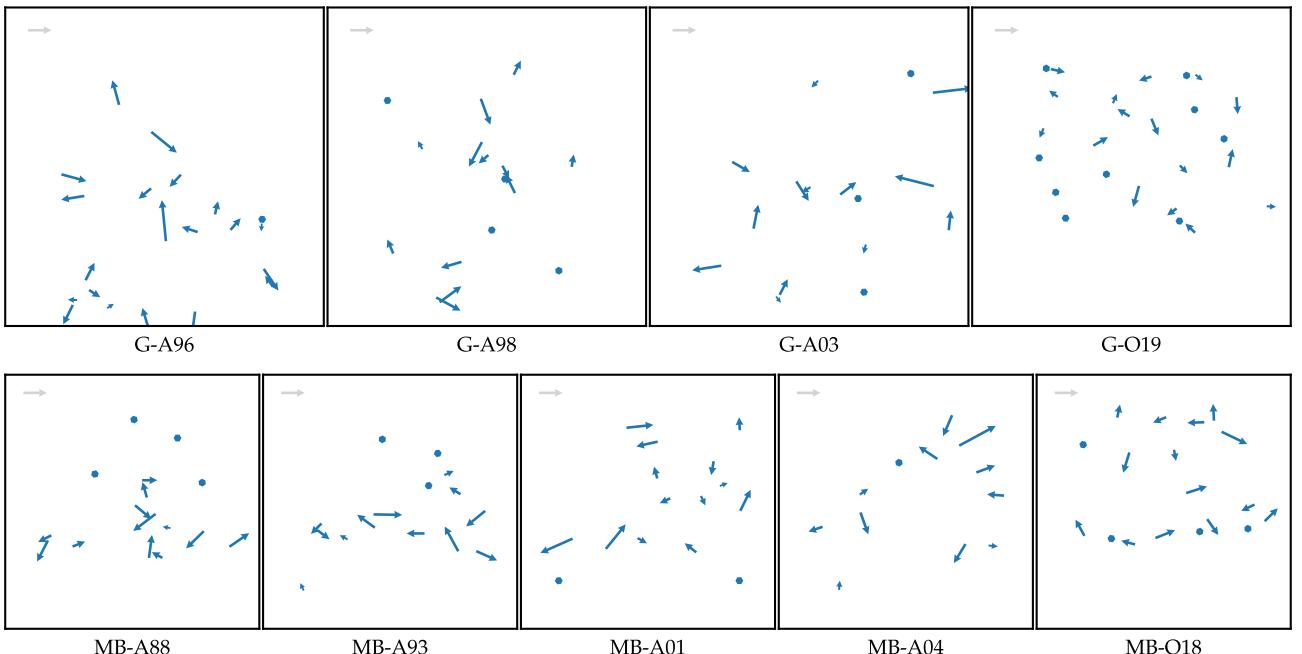


Fig. 10. 2D control point reprojection residual vectors in the image plane, shown for all fitted photographs. For each image, we derive reprojection error vectors normalized to control point uncertainties e_p , as described in Section 5.3.1. The resulting residuals are represented as dark blue vectors attached to their approximate respective control point location in the image plane. Control points with very small errors are represented as hexagonal dots. The light gray arrow in the upper left corner of each panel illustrates a unit length vector. The residual norms are consistent with the point placement uncertainties e_p , and the point estimate $\bar{\theta}$ provides a good fit to the control points. In addition, the residual vectors are fairly isotropic and exhibit no noticeable spatial trend.

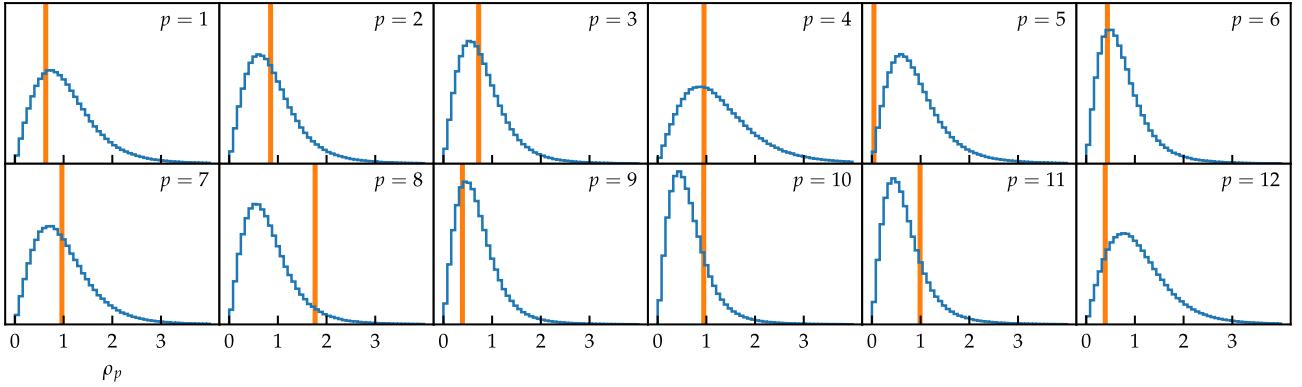


Fig. 11. Posterior predictive check on normalized residual norm for the control points of image A04. Each panel corresponds to a single 2D control point in this image. The horizontal axis corresponds to the norm of the (normalized) residuals ρ_p for each control point. For each point, the vertical bar marks ρ_p this point for the actual data D , while each histogram represents the posterior predictive distribution of ρ'_p under replicated observations D' predicted by the model based on inference from D . Please refer to the text of Section 5.3.2 for details. This test shows that the Bayesian model predicts replicated data D' which is compatible with the observed data D . The model appears self-consistent and does not exhibit any obvious tension with the data.

where in the first term on the right-hand side, we recognize the likelihood of D' (see Section 2.5), and the second term is the posterior of θ . We can therefore generate posterior predictive realizations of D' by sampling θ from the posterior $p(\theta|D, I)$ (i.e. taking the MCMC samples θ_i), and then for each sample θ_i , drawing a random realization $D' = \{(u'_p, v'_p)\}$ of 2D matches according to the likelihood (19).

As a test quantity for posterior predictive checks, we choose to look at the norm ρ_p of the normalized residual for control point p , as defined in Section 5.3.1. For the data $D = \{(u_p, v_p)\}$, this norm writes:

$$\rho_p := \frac{1}{e_p} \sqrt{(\tilde{u}_p(\bar{\theta}) - u_p)^2 + (\tilde{v}_p(\bar{\theta}) - v_p)^2}, \quad (35)$$

where like in Section 5.3.1, $\bar{\theta}$ designates the median point estimate for θ , and the tilde designates pixel coordinates for reprojected points. For replicated matches D' drawn from the posterior predictive distribution, the corresponding ρ'_p is defined identically, replacing the points (u_p, v_p) from D with (u'_p, v'_p) from D' . For a given image and set of matches D , we therefore obtain one single scalar value ρ_p of the test quantity for each control point p . The posterior predictive check amounts to comparing the value of ρ_p from the actual data D to the distribution of values ρ'_p when D' is sampled from the model's posterior predictive distribution.

Fig. 11 illustrates posterior predictive checking for image MB-A04. Each panel corresponds to one control point from this particular photograph. The normalized residual norm ρ_p for each point is represented along the horizontal axis. Vertical bars are located at ρ_p determined from the data, while the histograms show the posterior predictive distribution of the ρ'_p . We find that the distribution of ρ'_p is consistent with the measured ρ_p from the data, reproducing the magnitude and spread of the data well. We obtain similar results with the other images. This suggests that the model is able to self-consistently capture the statistical uncertainties in the control point placement process, and shows no sign of systematic discrepancies or tension.

5.4. Inverse perspective and surface area reconstruction

In this section, we test our method's ability to reconstruct the surface area of a polygonal feature drawn onto the 2D image. Here, we focus on the MB case study and its five photographs.

The polygonal region of interest is located on the Northwest face of the Mont Blanc massif, with a spatial extent of approximately 1400×500 meters, and chosen to present typical dimensions, shapes and slopes that could be encountered on a hanging glacier or an ice-covered rockwall, for example. Because the images for the MB case study were taken over a period of 30 years, we select the polygon contour based on features that are most likely fixed in the landscape; as

described in 5.1.1, we expect the topography of the chosen area to be relatively stable and therefore a good candidate for this validation test.

For each of the MB images, we trace a 2D polygon defined by the same terrain and geological features. The resulting polygons for all images are presented together in Fig. 12. The back projection relies on the posterior on camera parameters computed during the camera orientation and calibration process as described earlier.

In order to obtain an independent estimate of the surface area, we also carefully traced the same feature using the Google Earth Pro software. For all five images, we run the MCMC sampler as described in 4.1 using 2×10^4 MCMC steps and 400 walkers. For each image, its polygon is back projected 10^4 times as described in 4.2. We set $\sigma = 1$ pixel in the polygon tracing error model of Eq. 25. A random DEM is generated following 3.4 every 10 polygon back projection samples, to account for DEM-related uncertainties, with the parameter values for our DEM described in that section. For each image, this process yields an estimation of the posterior $p(S|D, \tilde{P}, \tilde{H}, I)$ based on 10^4 area samples.

The resulting posterior distributions for all images are presented in Fig. 13. Aerial images MB-A88, MB-A93, MB-A01 and MB-A04 use the distortion prior for aerial cameras. The terrestrial oblique photograph MB-O18 uses the broader general distortion prior. The vertical dashed line represents the value for the area obtained from Google Earth Pro. All posterior distributions are consistent with each other, as well as with the independent measurement of the area. Because the polygons are traced independently across all images, this shows that the method produces consistent outputs, but also that the systematic effects caused by manual tracing of the polygon are well captured in the variance of the posterior.

The posterior of image MB-O18 is much broader due to a combination of effects: the lower image resolution results in less accurate determination of both control points and traced polygonal outline, the weaker prior on the distortion introduces more statistical uncertainty, and finally, the oblique viewing angle makes the area more sensitive to back projection errors.

5.5. Impact of lens distortion

We now turn to the effect of lens distortions on area determination. For this test, we selected three images (MB-A01, MB-A04 and MB-O18) that we process first with their normal distortion priors (aerial priors for MB-A01 and MB-A04; and consumer lens for MB-O18). We then process each one again, but using the other prior: aerials are processed with the consumer lens prior which allows for distortion, producing the posteriors MB-A01 + D and MB-A04 + D (with “+D” for Distortion);

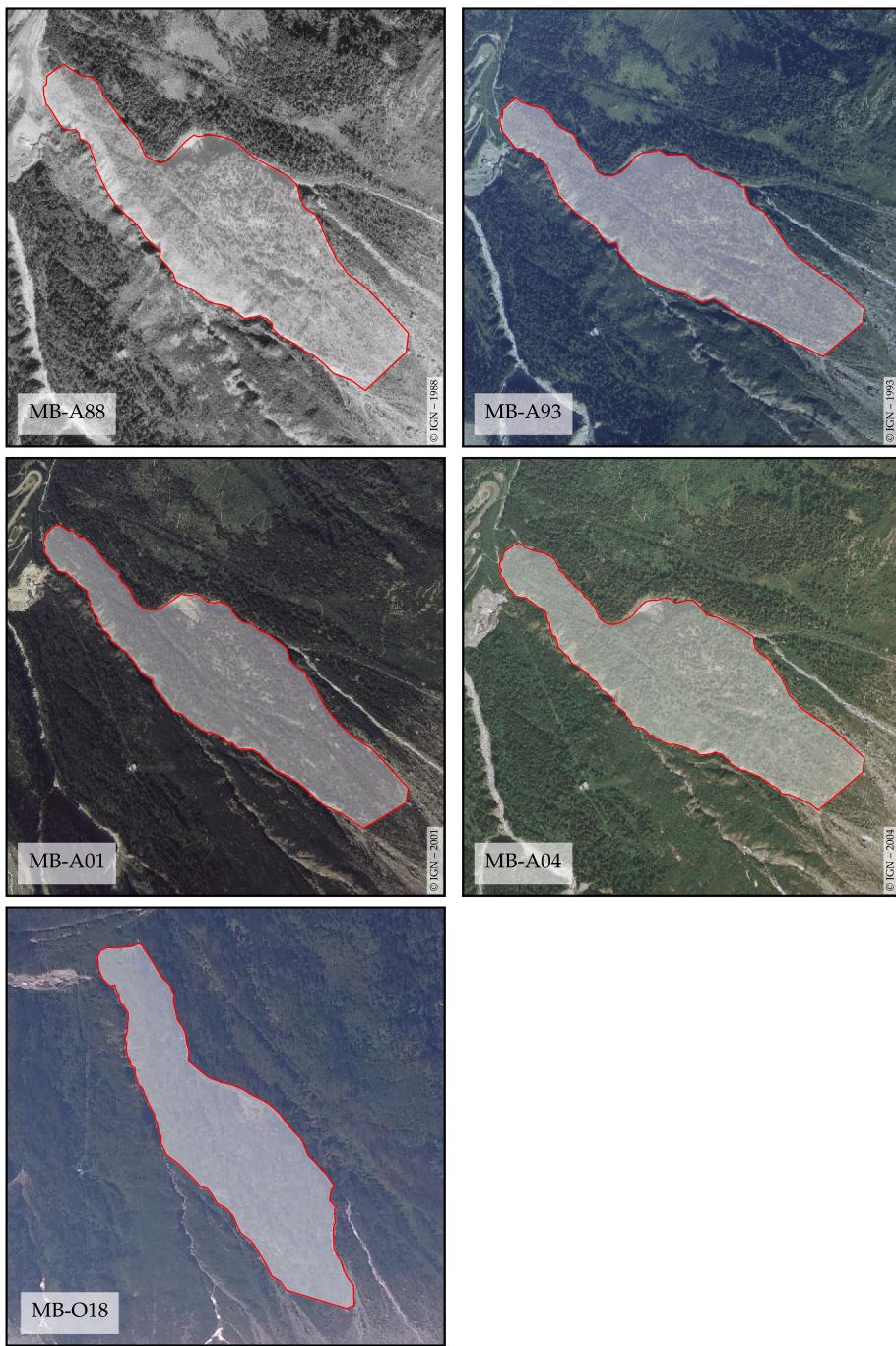


Fig. 12. 2D polygons traced on the test images of the MB case study, used for the validation of surface area reconstruction. The test polygons trace a same contour based on relatively fixed geological features, which are not expected to evolve over the 30 years spanned by the images.

while the terrestrial smartphone photograph is processed with the strict aerial prior, yielding the posterior MB-O18 + ND (“+ND” for No Distortion). The posteriors for all three pairs of models are shown in Fig. 14.

Aerial images MB-A01 and MB-A04 are insensitive to the introduction of lens distortions in the model. The posteriors of MB-A01 + D and MB-A04 + D are very slightly broader than MB-A01 and MB-A04 respectively, due to some additional statistical variance from the uncertainty on the distortion parameters. This shows that the distortion parameters are well constrained by control points for these high-quality images.

For the terrestrial oblique image MB-O18 taken using a smartphone, however, assuming low distortion results in a significant bias: the MB-

O18 + ND posterior excludes the actual value of the area by 4 standard deviations. When using the wider prior for consumer lenses (MB-O18), the bias is eliminated, at the cost of some additional statistical variance. This is an example of bias-variance trade-off, where a bias caused by systematic tension can be absorbed by a more general, better-fitting model, at the cost of some additional statistical uncertainty coming with the newly introduced parameters.

This test illustrates the importance of accounting for lens distortions when using photographs originating from unknown cameras, and shows that the model is able to correct them and eliminate bias, with very little consequence for aerial images which have low distortion in the first place.

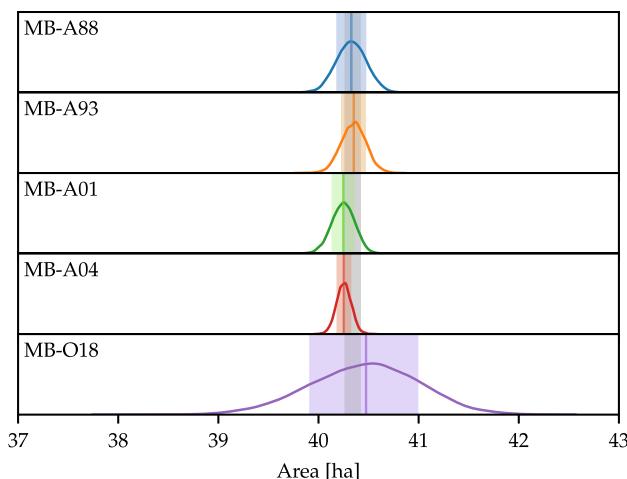


Fig. 13. Posterior distribution for the surface area of the polygon, for each test image. Aerial images MB-A88, MB-A93, MB-A01 and MB-A04 use the strict prior on distortion, while the terrestrial oblique photograph MB-O18 uses the weaker prior. The vertical gray span represents a value for the area independently derived using Google Earth Pro, covering a range of $\pm 0.2\%$ for reference. For each distribution, the median is represented using a vertical line, and the 68% credible region around the median (between the 16th and 84th percentiles) is shown by a shaded rectangle. All posterior distributions are consistent with each other, as well as with the independent measurement of the area, showing that statistical errors and systematic effects are well captured by the model.

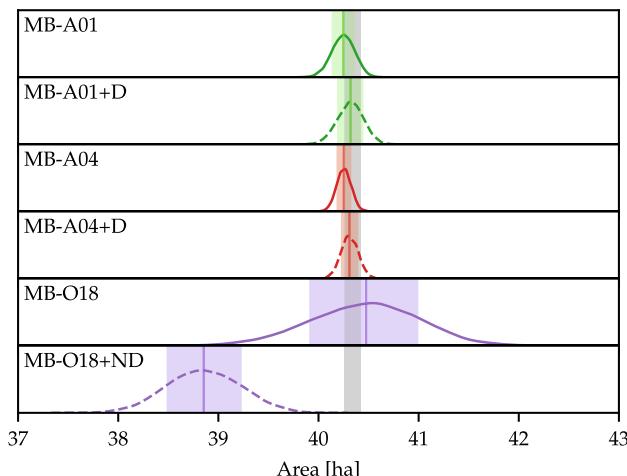


Fig. 14. Posterior distributions for the surface area of the polygon, illustrating the relevance of lens distortions for area back projection. Each pair (MB-A01, MB-A01 + D), (MB-A04, MB-A04 + D) and (MB-O18, MB-O18 + ND) corresponds to an image, processed with different distortion priors. Posteriors MB-A01 + D and MB-A04 + D allow for significant distortions on aerial images, which otherwise use the tighter aerial distortion prior; whereas posterior MB-O18 + ND uses the aerial prior for the low-grade smartphone lens, effectively enforcing a fit with very low distortions. For the aerial images MB-A01 and MB-A04, the impact of the lens distortion prior is not significant. For the terrestrial oblique image MB-O18, using the restrictive aerial prior results in a strong systematic bias of the estimated area, which disappears with the prior for consumer lenses.

6. Discussion

In Section 5.2 we compared the results of the camera calibration procedure to known values of camera parameters. We found that the errors of camera parameters reconstruction for terrestrial oblique pictures are of the same order of magnitude (a few meters) than the basic sources of uncertainty in the method such as the 3D position of the

ground control points. This indicates an accurate reconstruction of the camera position by our method, to the intrinsic level of accuracy allowed by the data and prior information.

For aerial photographs, we observe larger absolute reconstruction errors on the camera location, of the order of 50 m, despite higher image resolution. This discrepancy, however, is consistent in direction and magnitude with a sub-second delay between the last GPS fix information and the actual exposure. In any case, those discrepancies do not seem to be detrimental to the quality of the solution to the inverse perspective problem considered in this paper, as evidenced by the surface area test.

We showed in Section 5.4 that the presented method can successfully be used to estimate the surface area of landscape features on photographs. In Section 5.5, we have illustrated the need to account for possible lens distortions in the surface area estimation process. We found that, with a low-quality lens and without allowing for lens distortions in the model, the resulting estimated surface area is biased. We showed that this bias is eliminated, at the cost of additional variance, by relaxing the priors on distortion in the camera model. We find that allowing for distortion in the model does not impact surface area estimation derived from aerial photographs.

Our method accounts for uncertainties both in the manual tracing of the polygon, and in the digital elevation model used for back projection. We correspondingly proposed probabilistic models for these sources of errors, in Sections 3.3 and 3.4 respectively.

While providing an accurate description of user-generated uncertainties is a complicated task, results provided by back projecting the polygon using independent camera samples and our proposed error model are consistent with each other, and with independent measurements.

While we used this method to extract surface area measurements from photographic data, it can be extended to other types of measurements and inverse perspective problems. We also stress that the derivation of the posterior of S in Section 3.2 is very general, and can be extended in a straightforward fashion not only to other observables, but also to other potential sources of uncertainties, provided that they can be modeled satisfactorily.

7. Conclusion

In this paper, we presented a novel method for estimating surface area information from landscape features using single aerial and terrestrial photographs. Driven by the goal of characterizing uncertainties on the solutions of inverse perspective problems for archival or non-scientific photographs, we introduced models for errors in input data, as well as for characterizing uncertainties in digital elevation models. We integrated these ingredients into a statistically consistent Bayesian framework, which can readily be extended to other sources of uncertainty.

We applied and validated the method to camera orientation and calibration, before estimating the area of back projected polygons. Our method produced reconstructions of the camera parameters for aerial and terrestrial oblique photographs which are accurate to the level of our control points and input data. Combining Bayesian camera orientation and calibration together with probabilistic uncertainty models, the method consistently estimates surface areas from different photographs and camera angles, while also propagating the associated uncertainties. Applied to our particular problem, our study showed the potential importance of accounting for radial lens distortions when working with unknown consumer optics, as it can otherwise result in significant biases.

As future work, we intend to apply the method to estimate surface area variations in steep glaciers in the Mont Blanc massif, relying on a large set of terrestrial oblique and aerial images. The joint increase in availability of consumer-grade cameras (smartphones and small portable camcorders among others) and computing power provides the

scientific community with a large amount of quantitative physical landscape information to be extracted from single photographs. Bayesian methods as presented in this study stand out as valuable tools to build sophisticated models for inverse perspective and other photogrammetric methods, while providing an adaptable framework to estimate and control uncertainties in the results.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

The authors would like to thank the anonymous reviewers for their feedback and comments, which have contributed to significantly improving the manuscript. We also would like to thank the French Institut Géographique National (IGN) for helpful exchanges, and for providing us with valuable information on the aerial missions. Some map data is copyrighted by OpenStreetMap contributors and available from <https://www.openstreetmap.org>. Funding: This study is part of the ANR 14-CE03-0006 VIP Mont Blanc and the ALCOTRA AdaPT Mont Blanc projects.

References

- Bozzini, C., Conedera, M., Krebs, P., 2012. A new monoplanning tool to extract georeferenced vector data and orthorectified raster data from oblique non-metric photographs. *Int. J. Heritage Digital Era* 1 (3), 500–518. <https://doi.org/10.1260/2047-4970.1.3.499>.
- Brown, D.C., 1971. Close-range camera calibration. *Photogram. Eng.* 37 (8), 855–866.
- Carlisle, B.H., 2005. Modelling the spatial distribution of DEM error. *Trans. GIS* 9 (4), 521–540. <https://doi.org/10.1111/j.1467-9671.2005.00233.x>.
- Cekada, M.T., Zorn, M., Colucci, R.R., 2016. Monitoring Glacier Changes with the Use of Archive Images: The Example of the Julian Alps (NW Slovenia, NE Italy). In: Sustainable Mountain Regions: Challenges and Perspectives in Southeastern Europe. Springer, Cham, pp. 233–242. https://doi.org/10.1007/978-3-319-27905-3_17.
- Chapuis, A., Rolstad, C., Norland, R., 2010. Interpretation of amplitude data from a ground-based radar in combination with terrestrial photogrammetry and visual observations for calving monitoring of Kronebreen, Svalbard. *Annals Glaciol.* 51 (55), 34–40. <https://doi.org/10.3189/172756410791392781>.
- CIPA Standardization Committee, Guideline for Noting Digital Camera Specifications in Catalogs, Revised Version (Oct. 2005).
- Criminisi, A., 2001. Accurate Visual Metrology from Single and Multiple Uncalibrated Images, Distinguished Dissertations. Springer-Verlag, London.
- European GNSS Agency, EGNOS Open Service (OS) Service Definition Document (Oct. 2017).
- Feurer, D., Vinatier, F., 2018. Joining multi-epoch archival aerial images in a single SfM block allows 3-D change detection with almost exclusively image information. *ISPRS J. Photogram. Remote Sens.* 146, 495–506. <https://doi.org/10.1016/j.isprsjprs.2018.10.016>.
- Fisher, P., 1991. First Experiments in Viewshed Uncertainty: The Accuracy of the Viewshed Area, Photogrammetric Engineering & Remote Sensing 57.
- Fisher, P., 1998. Improved modeling of elevation error with geostatistics. *GeoInformatica* 2, 215–233. <https://doi.org/10.1023/A:1009717704255>.
- Fonseca, T.C.O., Ferreira, M.A.R., Migon, H.S., 2008. Objective Bayesian analysis for the Student-t regression model. *Biometrika* 95 (2), 325–333. <https://doi.org/10.1093/biomet/asn001>.
- Foreman-Mackey, D., Hogg, D.W., Lang, D., Goodman, J., 2013. Emcee: the MCMC Hammer. *Publ. Astron. Soc. Pac.* 125 (925), 306–312. <https://doi.org/10.1086/670067>. arXiv:1202.3665.
- Förstner, W., Wrobel, B.P., 2016. Photogrammetric Computer Vision, Vol. 11 of Geometry and Computing, Springer International Publishing, Cham. <https://doi.org/10.1007/978-3-319-11550-4>.
- Gelman, A., Meng, X.-L., Stern, H., 1996. Posterior predictive assessment of model fitness via realized discrepancies. *Stat. Sin.* 6 (4), 733–760.
- Gelman, A., Carlin, J.B., Stern, H.S., Dunson, D.B., Vehtari, A., Rubin, D.B., 2013. Bayesian Data Analysis, 3rd ed. CRC Press.
- Ghosh, S.K., 2005. Fundamentals of Computational Photogrammetry. Concept Publishing Company.
- Goodman, J., Weare, J., 2010. Ensemble samplers with affine invariance. *Commun. Appl. Math. Comput. Sci.* 5 (1), 65–80. <https://doi.org/10.2140/camcos.2010.5.65>.
- Gruen, A., 1985. Adaptive least squares correlation: a powerful image matching technique. *South African J. Photogram., Remote Sens., Cartogr.* 14, 175–187.
- Gruen, A., Li, H., 1995. Road extraction from aerial and satellite images by dynamic programming. *ISPRS J. Photogram. Remote Sens.* 50 (4), 11–20. [https://doi.org/10.1016/0924-2716\(95\)98233-P](https://doi.org/10.1016/0924-2716(95)98233-P).
- Haala, N., Brenner, C., 1999. Extraction of buildings and trees in urban environments. *ISPRS J. Photogram. Remote Sens.* 54 (2), 130–137. [https://doi.org/10.1016/S0924-2716\(99\)00010-6](https://doi.org/10.1016/S0924-2716(99)00010-6).
- Haneberg, W., 2006. Effects of digital elevation model errors on spatially distributed seismic slope stability calculations: an example from Seattle, Washington. *Environ. Eng. Geosci.* 12, 247–260. <https://doi.org/10.2113/gsgeosci.12.3.247>.
- Heikkilä, J., Silven, O., 1997. A four-step camera calibration procedure with implicit image correction. In: Proceedings of IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 1997, pp. 1106–1112. <https://doi.org/10.1109/CVPR.1997.609468>.
- Heipke, C., 1992. A global approach for least-squares image matching and surface reconstruction in object space. *Photogram. Eng.* 7.
- Hobbie, D., 2010. The development of photogrammetric instruments and methods at Carl Zeiss in Oberkochen, 2010.
- Hogg, D.W., Foreman-Mackey, D., 2017. Data analysis recipes: Using Markov Chain Monte Carlo, arXiv:1710.06068 [astro-ph, physics:physics, stat]arXiv:1710.06068.
- Hogg, D.W., Bovy, J., Lang, D., 2010. Data analysis recipes: Fitting a model to data, arXiv:1008.4686 [astro-ph, physics:physics]arXiv:1008.4686.
- Holmes, K.W., Chadwick, O.A., Kyriakidis, P.C., 2000. Error in a USGS 30-meter digital elevation model and its impact on terrain modeling. *J. Hydrol.* 233 (1), 154–173. [https://doi.org/10.1016/S0022-1694\(00\)00229-8](https://doi.org/10.1016/S0022-1694(00)00229-8).
- Hunter, G.J., Goodchild, M.F., 1997. Modeling the uncertainty of slope and aspect estimates derived from spatial databases. *Geogr. Anal.* 29 (1), 35–49. <https://doi.org/10.1111/j.1538-4632.1997.tb00944.x>.
- I.G. National, Remonter le temps, <https://remonterletemps.ign.fr/>, Jul. 2019.
- Jauregui, M., Vilchez, J., Chacón, L., 2002. A procedure for map updating using digital mono-plotting. *Comput. Geosci.* 28 (4), 513–523. [https://doi.org/10.1016/S0098-3004\(01\)00068-1](https://doi.org/10.1016/S0098-3004(01)00068-1).
- Jordan, E., Ungerechts, L., Cáceres, B., Peñafiel, A., Francou, B., 2005. Estimation by photogrammetry of the glacier recession on the Cotopaxi Volcano (Ecuador) between 1956 and 1997. *Hydrog. Sci. J.* 50 (6), 949–961. <https://doi.org/10.1623/hysj.2005.50.6.949>.
- Kannala, J., Brandt, S.S., 2006. A generic camera model and calibration method for conventional, wide-angle, and fish-eye lenses. *IEEE Trans. Pattern Anal. Mach. Intell.* 28 (8), 1335–1340. <https://doi.org/10.1109/TPAMI.2006.153>.
- Krahnstöver, N., Mendonça, P.R.S., 2005. Bayesian autocalibration for surveillance, in: Tenth IEEE International Conference on Computer Vision (ICCV'05) Volume 1, Vol. 2, 2005, pp. 1858–1865. <https://doi.org/10.1109/ICCV.2005.44>.
- Küng, O., Strecha, C., Beyerle, A., Zufferey, J.-C., Floreano, D., Fua, P., Gervaux, F., 2012. The accuracy of automatic photogrammetric techniques on ultra-light UAV imagery. *Int. Arch. Photogram., Remote Sens. Spatial Inform. Sci. – ISPRS Arch.* 38, 125–130. <https://doi.org/10.5194/isprarchives-XXXVIII-1-C22-125-2011>.
- Kyriakidis, P.C., Shortridge, A.M., Goodchild, M.F., 1999. Geostatistics for conflation and accuracy assessment of digital elevation models. *Int. J. Geogr. Inform. Sci.* 13 (7), 677–707. <https://doi.org/10.1080/136588199241067>.
- Lapine, L.A., 1996. Airborne Kinematic GPS Positioning for Photogrammetry - The Determination of the Camera Exposure Station, Tech. rep., MD, Silver Springs.
- Lee, J., 1992. Modeling the effect of data errors on feature extraction from digital elevation models. *Photogram. Eng.* 7.
- Li, Z., 1988. On the measure of digital terrain model accuracy. *Photogram. Rec.* 12 (72), 873–877. <https://doi.org/10.1111/j.1477-9730.1988.tb00636.x>.
- Lichti, D., 2002. The interpolation problem in GPS-supported aerial triangulation. *Photogram. Rec.* 17 (99), 481–492. <https://doi.org/10.1111/0031-868X.00200>.
- Liu, H., Jezek, K.C., 1999. Investigating DEM error patterns by directional variograms and Fourier analysis. *Geogr. Anal.* 31 (3), 249–266. <https://doi.org/10.1111/j.1538-4632.1999.tb00981.x>.
- Micheletti, N., Lane, S.N., Chandler, J.H., 2015. Application of archival aerial photogrammetry to quantify climate forcing of alpine landscapes. *Photogram. Rec.* 30 (150), 143–165. <https://doi.org/10.1111/phor.12099>.
- Murtiyoso, A., Remondino, F., Rupnik, E., Nex, F., Grussenmeyer, P., 2014. Oblique aerial photography tool for building inspection and damage assessment. *ISPRS - Int. Arch. Photogram., Remote Sens. Spatial Inf. Sci.* XL-1, 309–313. <https://doi.org/10.5194/isprarchives-XL-1-309-2014>.
- Olson, L.N., Schroder, J.B., 2018. PyAMG: Algebraic Multigrid Solvers in Python v4.0, release 4.0.
- OpenStreetMap contributors, 2019. Planet dump retrieved from <<https://planet.osm.org>>.
- Piermattei, L., Carturan, L., Guarneri, A., 2015. Use of terrestrial photogrammetry based on structure-from-motion for mass balance estimation of a small glacier in the Italian alps. *Earth Surf. Proc. Land.* 40 (13), 1791–1802. <https://doi.org/10.1002/esp.3756>.
- Produit, T., Ingensand, J., 2016. A 3D georeferencer and viewer to relate landscape pictures with VGI, in: AGILE 2016, Helsinki, 2016, p. 4.
- RTM ONF, Plan de Prévention des Risques naturels prévisibles, carte des aléas, Chamonix, Apr. 2002.
- RTM ONF, Base de données RTM, 2019.
- Scapozza, C., Schenker, F.L., Castelletti, C., Bozzini, C., Ambrosi, C., 2016. Digital mono- and 3D stereo-photogrammetry for geological and geomorphological mapping. In: EGU General Assembly Conference Abstracts, Vol. 18, pp. EPSC2016-4787.
- Shah, A., Wilson, A.G., Ghahramani, Z., 2014. Student-t Processes as Alternatives to Gaussian Processes, arXiv:1402.4306 [cs, stat] (2014) 877–885arXiv:1402.4306.
- Sheng, Y., 2005. Theoretical analysis of the iterative photogrammetric method to determine ground coordinates from photo coordinates and a DEM. *Photogram. Eng. Remote Sens.* 71 (7), 863–871. <https://doi.org/10.14358/PERS.71.7.863>.
- Sheynin, O., 1995. Helmert's work in the theory of errors. *Arch. Hist. Exact Sci.* 49 (1), 73–104. <https://doi.org/10.1007/BF00374700>.

- Sivia, D., Skilling, J., 2006. *Data Analysis: A Bayesian Tutorial*, second ed. Oxford University Press, Oxford, New York.
- Stockdale, C.A., Bozzini, C., Macdonald, S.E., Higgs, E., 2015. Extracting ecological information from oblique angle terrestrial landscape photographs: performance evaluation of the WSL Monoplotting Tool. *Appl. Geogr.* 63, 315–325. <https://doi.org/10.1016/j.apgeog.2015.07.012>.
- Strausz, D.A., 2001. Application of Photogrammetric Techniques to the Measurement of Historic Photographs, Tech. rep., Oregon State University, Department of Geosciences.
- Streilein, A., 1994. Towards automation in architectural photogrammetry: CAD-based 3D-feature extraction. *ISPRS J. Photogram. Remote Sens.* 49 (5), 4–15. [https://doi.org/10.1016/0924-2716\(94\)90019-1](https://doi.org/10.1016/0924-2716(94)90019-1).
- Sturm, P., 2014. Pinhole Camera Model, in: Computer Vision, Springer, Boston, MA, 2014, pp. 610–613. https://doi.org/10.1007/978-0-387-31439-6_472.
- Sturm, P., Ramalingam, S., Tardif, J.-P., Gasparini, S., Barreto, J., 2011. Camera models and fundamental concepts used in geometric computer vision. *Found. Trends Comput. Graph. Vision* 6 (1–2), 1–183. <https://doi.org/10.1561/0600000023>.
- Sundareswara, R., Schrater, P., 2005. Bayesian Modelling of Camera Calibration and Reconstruction. In: Fifth International Conference on 3-D Digital Imaging and Modeling (3DIM'05). IEEE, Ottawa, ON, Canada, pp. 394–401. <https://doi.org/10.1109/3DIM.2005.24>.
- Tayman, W.P., 1984. User Guide for the USGS aerial camera report of calibration. *Photogram. Eng. Remote Sens.* 50 (5), 577–584.
- US Geological Survey, USGS Aerial Camera Specifications, Jul. 2012.
- Valkenburg, R.J., 1998. A Bayesian Approach to Camera System Calibration/Spatial Intersection. In: IVCNZ, Citeseer, pp. 11–16.
- Wang, J., Shi, F., Zhang, J., Liu, Y., 2008. A new calibration model of camera lens distortion. *Pattern Recogn.* 41 (2), 607–615. <https://doi.org/10.1016/j.patcog.2007.06.012>.
- Warner, W.S., Andersen, S., Saeland, S., 1993. Surveying a waste site with 35-mm oblique aerial photography: monoplotting with a digitizing tablet. *Cartogr. Geogr. Inform. Syst.* 20 (4), 237–243. <https://doi.org/10.1559/152304093782637505>.
- Wechsler, S.P., Kroll, C.N., 2006. Quantifying DEM uncertainty and its effect on topographic parameters. *Photogram. Eng. Remote Sens.* 72 (9), 1081–1090. <https://doi.org/10.14358/PERS.72.9.1081>.
- Weng, J., Cohen, P., Herniou, M., 1992. Camera calibration with distortion models and accuracy evaluation. *IEEE Trans. Pattern Anal. Mach. Intell.* 14 (10), 965–980. <https://doi.org/10.1109/34.159901>.
- Wiesmann, S., Steiner, L., Pozzi, M., Bozzini, C., Bauder, A., Hurni, L., 2012. Reconstructing Historic Glacier States Based on Terrestrial Oblique Photographs. In: Proceedings - AutoCarto 2012 - Columbus, Ohio, USA - September 16–18, 2012, Cartography and Geographic Information Society (CaGIS), 2012, p. 14.
- Willneff, J., Poon, J., Fraser, C., 2005. Monoplotting applied to high-resolution satellite imagery. *J. Spatial Sci.* 50 (2), 1–11. <https://doi.org/10.1080/14498596.2005.9635045>.
- Zhang, Z., 2000. A Flexible New Technique for Camera Calibration, *IEEE Trans. Pattern Anal. Mach. Intell.* 22.