Impact of image quality on SfM Photogrammetry: colour, compression and noise

James O'Connor

This thesis is being submitted in partial fulfilment of the requirements of Kingston University for Doctor of Philosophy

Collaborating establishment: Lancaster University

Date of submission: 28/09/2018

Acknowledgements

Firstly, I would like to thank Mike Smith, my supervisor for most of the duration of my PhD candidacy. His immense patience, guidance and very positive attitude towards work has allowed for me to develop into a better scientist. He has been incredibly supportive of this work and the direction I took the project over its course and has made me a much more passionate photographer to boot! I look forward to working with him in the future.

Secondly, I would like to thank Mike James, my co-supervisor, who has been very good at keeping me up-to-date with the latest trends in photogrammetry. The many discussions we had over accuracy within photogrammetry have nurtured a love of precision and have allowed me to think more critically and objectively about research within the geoscientific community. His talk at the RSPSoc annual conference at Imperial College in 2017 was incredibly inspirational and provided me with the drive needed to finish this thesis.

Nigel Walford was great within the last year of this project, during the writing of this work. His support was very much appreciated! I look forward to visiting him at Kingston in the coming years.

My family have been amazing at encouraging me to do research I'm passionate about and try and make a positive impact in this world. I would like to convey a big thanks to them.

Contents

Glossary of terms

Acronym	Definition
C2C	Cloud-to-cloud distance
DEM	Digital elevation model (Surface with vegetation/man-made features removed – 'bare earth')
DSM	Digital surface model (Surface with vegetation/man-made features included)
DoD	Difference of DSM
FOV	Field of view
GCP	Ground Control Point
GPS-IMU	Global positioning system-Inertial measurement unit
GSD	Ground sample distance
ICP	Iterative closest point
ILC	Interchangeable lens camera
M3C2	Multiscale model to model cloud comparison
ME	Median Error
PCA	Principle components analysis
RANSAC	Random sampling consensus
RMSE-P	Root mean square pixel matching error
SfM	Structure-from-motion
SNR	Signal to noise ratio
TLS	Terrestrial Laser Scanner
TS	Total Station
UAV	Unmanned aerial vehicle

Abstract

Structure-from-motion (SfM) photogrammetry has become ubiquitous in the geosciences, owing to its low-cost and ease of use for generating 3D data. Ideas around data collection, quality and processing need to be revisited to ensure that the technology is being harnessed correctly. One area which is new in this current image acquisition boom is the range of sensors and systems being used to collect image data. This raises crucial questions in the geoscience community which are addressed in this contribution. This is split into three parts.

Firstly, image quality is investigated to establish whether a stable association between it and the quality of photogrammetric products can be uncovered to allow simpler and more effective inter-comparison of results between studies. This was accomplished by artificially degrading a very high-quality benchmark dataset of a coastal cliff and a landslide in Norfolk, UK. Results revealed that the level of noise, image compression and downsampling all degrade the quality of 3D products from the SfM workflow.

Secondly, these sets of images were pre-processed to establish whether results could be augmented by controlling the single colour channel used during photogrammetric processing. Results showed slight variations in the products generated, with evidence supporting the fine sensitivity SfM has for refining the focal length estimation of the lens. For extremely specific contexts, pre-processing of the RGB-to-single channel conversion may be relevant, but for the datasets analysed in this contribution this was not the case.

Thirdly, image network configurations were investigated to build on previous research in establishing best practice. Results show that, in situations where the number of images being acquired is a limiting factor, networks with narrowly oblique overlapping images have a higher density and lower error than those with widely oblique images and those directly facing the surface normal.

These results demonstrate the value of optimising image acquisition, and in the handling of this imagery. The differences in image quality and pre-processing which are unreported within geoscientific studies using SfM could account for differences between accuracies obtained, independent of the specific photogrammetric methods used. Insights from this work into how best to capture, process and produce the best quality SfM data will allow the community to adopt these best practices in the future.

Chapter 1. Introduction

Photogrammetry, the science of measuring distances from photographs, dates from the middle of the 19th century, with the first technology emerging for acquiring and developing film photographs. With the advents of aerial platforms for photography and later digital photography, photogrammetry's utility in the geosciences has grown and it is now commonly used as a primary source for 3D geographic data.

Over the last two decades, photogrammetry has changed dramatically in how frequently it is used within geoscience, owing to the low barriers to entry which currently exist. These low barriers are a result of rapid software development within a field of photogrammetry known as 'structure-from-motion' (SfM). SfM photogrammetry requires almost no knowledge of image acquisition or the photogrammetric workflow to produce 3D data, and so has proved popular amongst amateurs, researchers and professionals alike.

1.1 Applications of photogrammetry in the geosciences

Digital cameras and smart phones containing cameras have been used extensively for photogrammetry in the geosciences, from fields as wide as landform monitoring of active volcanos to the retrieval of small scale forest structure (Nakano et al., 2014; Puliti et al., 2015). Traditionally, a metric camera would be required to acquire data for use in photogrammetry, which had an associated camera model, correcting for distortions within the camera system (Lane et al., 2000). These images would be used within specialist photogrammetric software packages to generate digital surface models (DSMs) which could be used within geoscientific studies.

SfM photogrammetry simultaneously removed both the barrier of needing a high-quality metric camera and expertise in the use of photogrammetric software, allowing geoscientists to rapidly collect 3D data with little training required. It allows for production of 3D data from unordered image collections, with no prior knowledge of camera positions/metadata (Snavely et al., 2008). SfM software is widely available, including open-source solutions (VisualSFM, MicMac). Thus, it has opened photogrammetry up to geoscientists with little prior experience of photogrammetry. This is evidenced by the amount of published literature within geoscientific journals mentioning photogrammetry each year, according to Web of Science (Figure 1.1).

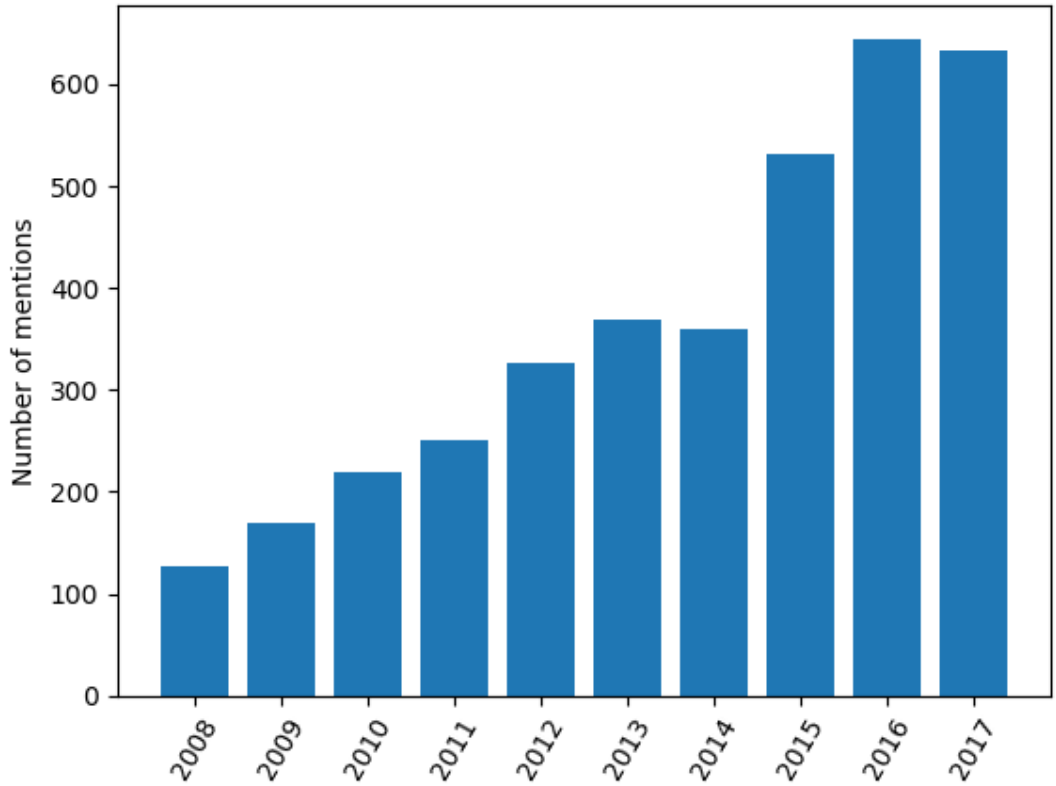

Figure 1.1. Number of journal publications referring to "photogrammetry" within the 'Remote Sensing', 'Geosciences Multidisciplinary' and 'Geography Physical' categories on Web of Science (Web of Science, 2018).

With this growth in use amongst researchers from outside the photogrammetric community comes a range of experience in image acquisition. This work seeks to establish how results from similar pieces of photogrammetric work vary given a range of image quality, and how correct reporting of data can standardise image quality reporting within the geoscientific community. It analyses aspects of photogrammetry, image quality and their interaction to understand this issue from a geoscientific context.

1.2 Image quality

Image quality is a general term used to describe attributes within a given image, whether analog or digital. Different imaging configurations (camera and lens combinations) will produce images of different quality. The main factors which affect image quality are summarised in Table 1.1

Table 1.1 Main factors influencing image quality

Factor	Summary
Noise	Random changes in brightness values due to variable photon arrival rate and imperfections in camera electronics.
Compression	Compression of image data to reduce file sizes.
Sharpness	How well contrast is represented in an image versus reality
Dynamic range	The range of brightness values which can be retrieved within an image.
Distortion	Curving of straight lines due to imperfections in lens design.
Bit depth	The number of bits a single piece of data are written to

Noise will ensure that two images taken from the same imaging configuration will almost never be identical, but these noise levels are generally predictable, so users can expect a certain noise level for a sensor prior to acquiring imagery. Image compression can introduce irreversible loss of data to the imagery acquired, depending on the algorithm chosen. Compression is common when acquiring imagery and is often applied automatically by the image processing software onboard the camera and users may not be able to access the raw image data. Sharpness of imagery varies depending on the lens being used for image acquisition. Generally, low quality lenses will have lower sharpness, and this sharpness will drop off towards the edges of images. Dynamic range will principally affect images which require a large range of brightness values to be covered, and so may not play a factor in image quality, depending on the contents of the image. Lastly, distortion is caused by lens defects, and is poor in particularly long or short focal length lenses.

Image quality depends on the camera and lens used to capture a given image, as well as the settings used for capturing that image. These have direct impacts on each of the factors listed in Table 1.1. Within the geosciences, image quality varies considerably between studies. Standard metrics for measuring image quality are often technical and laborious to generate, which likely contributes to the lack of their adoption within the scientific community outside of specialist image quality branches of research.

1.3 Gap in knowledge

There is limited research detailing the effects of varying image quality within photogrammetric research (some aspects are included within Middlebury, (2015)), and none specifically for geoscientific studies. This is important, as researchers need guidelines on the accuracies they can reasonably expect to attain based on previous literature. This gap in knowledge presents a

difficulty in comparison between research studies, research reproducibility and the validation of results across studies using different equipment, or the same equipment with different settings.

Research in the fields of image quality and photogrammetry are not new, but the interaction between these two fields, especially with the advent of SfM photogrammetry, is understudied. Thus, this work aims to clarify what these effects are, their scale, and how the community can best report these within the scientific literature.

In addition, while image processing for photogrammetry has been investigated in several scientific texts (Gómez-Gutiérrez et al., 2014; Verhoeven et al., 2015), the potential benefits of pre-processing images through to the final product of SfM photogrammetry remain understudied.

1.4 Summary of research goals

With this gap in knowledge highlighted above, this research targets aspects of image quality and image processing, and their relationship to photogrammetry. Specifically, the core aim is to test whether image quality and processing have a significant impact on the outputs from the SfM photogrammetric process.

To address whether their magnitude is measurable, images were acquired at two different locations. This allowed a range of experiments to be undertaken that varied image quality and image processing up to the stage of ingestion into the photogrammetric workflows, which enabled like-for-like comparison of each product within the test suite.

1.5 Thesis outline

With the scope of research outlined, Chapter 2 outlines and reviews current literature, introducing a relevant background to the topics introduced in Chapter 1. Chapter 3 features site descriptions, research design and the methods for data capture used. Chapter 4 focuses on image degradation and how this varies the quality of the output 3D data from SfM. Chapter 5 focuses on digital image processing and single channel generation using the image data acquired, and optimisation for use in the SfM processing workflow. Chapter 6 will focus on the impact of survey network design and how introducing more cameras affects the output accuracies of photogrammetric products. Chapter 7 presents common usage scenarios for SfM, and how the results of this research will affect how a survey should be approached. Chapter 8 discusses the research outcomes, detailing the main insights gained within this thesis. Lastly, chapter 9 will set out the conclusions to be drawn from this work, and future work which would develop this work further.

Chapter 2: Literature review

Within this chapter, the basic ideas of photogrammetry will be introduced, and literature will be discussed demonstrating the usefulness of photogrammetry as a tool in the geosciences by exemplifying a few of the research areas within which it is used. Next, a thorough review of each step involved in the photogrammetric workflow will be introduced, highlighting areas within the workflow this novel research seeks to optimize.

2.1 Principles of SfM photogrammetry

Photogrammetry uses information gained from parallax, the difference in apparent position of an object based on perspective, to establish 3D structure from a series of 2D images. Traditional photogrammetry is the term generally used to define photogrammetry where estimation of relative camera positions and calibration of a camera's internal geometry is not required. These surveys generally use a metric camera which has been lab-calibrated. A high-quality positioning device is used to recover very precise position and orientation data relating to the camera at time of exposure. This information is used in photogrammetric software to recreate the 3D object.

SfM photogrammetry refers to a novel group of algorithms which can act in conjunction with one another to provide topographic information from a set of unregistered images from an uncalibrated camera. SfM photogrammetry, this involves several stages of planning, provisioning and processing, as outlined in Figure 2.1.

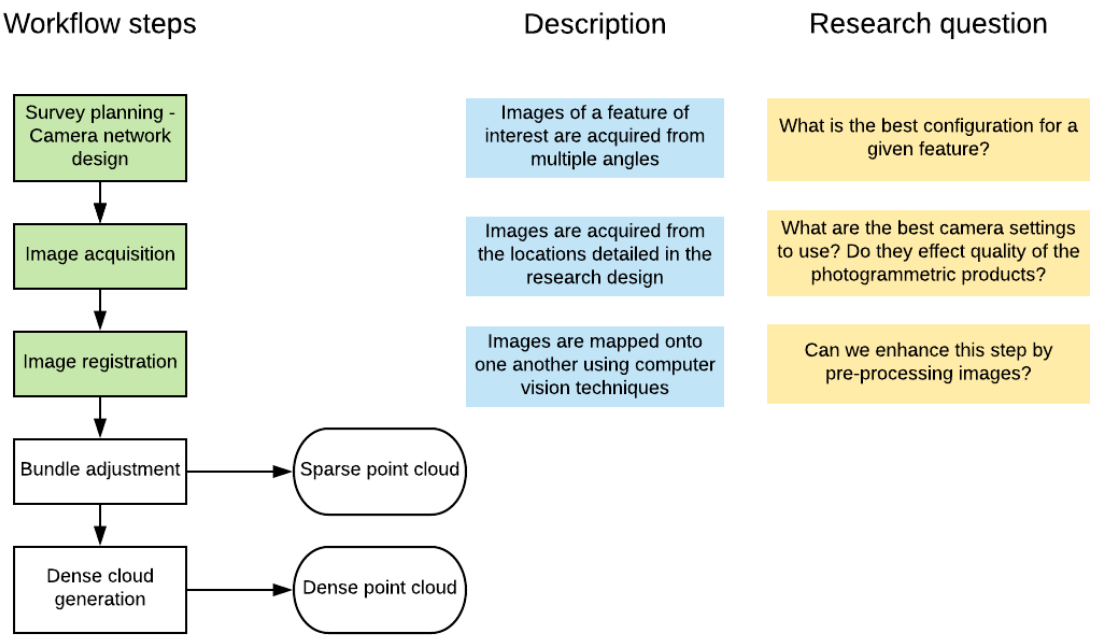

Figure 2.1. Workflow detailing main steps in the photogrammetric workflow, with the steps this research intends to fill gaps in knowledge highlighted in green.

In the first instance, a user will acquire a collection of images of the area of interest, moving the camera's position relative to the area to generate depth information ('motion' in Structure-from-motion). After acquisition, images are 'registered' onto one another. This involves mapping each image onto each other image within the set of images with which it shares common image information ('overlaps' with that image). The output of this process is a list of mappings of each image onto each other image within the image set.

There will be residual error when registering image features onto one another, which is leveraged within the next step, bundle adjustment, which refines the solution by optimising these residuals, as well as parameters within a model which represents the camera(s) from which data was acquired. After bundle adjustment, a 3D point cloud of the data is available for viewing, which represents the common points identified within the image set, as well as their locations in an arbitrary common coordinate system (a 'sparse' point cloud).

This is a stochastic process due to the non-linearity of the solver used. Thus, for an input series of images with a given set of photogrammetric parameters, the same sparse cloud won't be guaranteed each time the process is executed.

A further step, dense cloud generation, is available, which takes the input images, along with the output from the bundle adjustment, to densify the point cloud using the image information (resulting in a 'dense' point cloud, Hirschmüller (2008)).

2.2 Photogrammetry in the geosciences

The use of 3D data has seen rapid growth in the geosciences, and much of the literature comprises of researchers leveraging photogrammetry as a primary data source within geomorphological studies. It principally comes in two forms, Digital Surface Models (DSMs) are a continuous data which represent the topography of a surface. Point clouds are a discrete data, which represent measurements recorded at specific point positions across a surface or a volume. Within structure-from-motion (SfM) photogrammetry, point clouds are generally produced in the first instance, and DSMs derived from these if necessary. Typical activities/research involving SfM-derived 3D data include accuracy assessment against a higher quality ground truth (Micheletti et al., 2015) and change detection through time series analysis (Ryan et al., 2015; Williams et al., 2014). Many of these studies utilised Unmanned Aerial Vehicles (UAVs) which are mounted with cameras of varying quality. Owing to their use in differing environmental conditions, the settings used are highly variable also.

2.2.1 Forestry

Several studies have shown SfM photogrammetry, both aerial and ground based, to be highly effective in generating estimates of both stand size and canopy height (Bauwens et al., 2016; Mikita et al., 2016; Mlambo et al., 2017; Wallace et al., 2016). Mlambo et al., (2017) have shown aerial photogrammetry to be useful in recovering canopy height models in areas where the canopy is not closed (with pixel-wise correlations of derived DEMs showing an R^2 value of 0.89), though this completely changes in areas with closed canopy forests (the R^2 degrading to 0.17). This makes photogrammetry a promising technology for doing rapid volumetric surveys of woody plant matter during times of the year when leaves are off. Within this study, the camera chosen was a GoPro Hero 3+ with a very short focal length, relatively small sensor and wide-angled lens. Mikita et al., (2016) showed that using ground based photogrammetry, walking between tracks in forestry, forest stand volume could be measured. They conclude one can reasonably expect diameter-at-breast-height (DBH) to be within 1 cm of manual measurement, and tree height measurements to be within 1 m from the aerial imagery for spruce forestry. Their chosen camera was a Sony NEX5, which likely produced much higher quality images than that of the GoPro Hero within Mlambo's work, though this is currently unaccounted for within these contributions.

Bauwens et al., (2016) demonstrated the utility for measuring the DBH of irregularly shaped tree trunks in the tropics, which could reliably be used to estimate total biomass of a stand using typical allometric relationships which relate the size of the trunk to the overall biomass of a given tree.

These examples show SfM photogrammetry can be used in complex scenarios within the geosciences, but without proper error estimation and tracking, the quality of outputs are questionable as a result. Within the literature it is seen that vegetated areas are frequently problematic when it comes to producing accurate results, which remains a limitation of the technology when applied in these fields.

2.2.2 Landslides

Photogrammetry has been successfully applied to monitoring both short and long term change detection of landslides in many different geomorphological contexts (James et al., 2017a; Lucieer et al., 2014; Niethammer et al., 2010; Stumpf et al., 2015; Turner et al., 2015). Niethammer et al., (2010) used a UAV with a consumer off the shelf (COTS) camera to measure displacement against a high quality, high resolution orthophoto (geometrically corrected image where scale is uniform) acquired a year earlier. Worth noting is the lack of documentation of the specific camera used, and lack of rigour in reporting the settings used on the camera, highlighting the difficulty there would be to reproduce this research effectively. Stumpf et al., (2015) took multi-temporal measurements of a well-studied landslide (Super-Sauze), while simultaneously collecting terrestrial laser scan data to use as a reference. The study site covers an area approximately 900 m in length by 135 m in width. The images used within their study were acquired from many different and not-complimentary positions, however, with panoramic images being taken at very different ranges using two different protocols; one a range between 20 m – 200 m, and another between 50 m – 1000 m. Point cloud products generated using strictly open source tools (VisualSfM and Micmac) led to accuracies of within 5 cm. A novel algorithm, Multiscale Model to Model Cloud Comparison (Lague et al., 2013), was used to track displacements in the scene. These studies set a high quality baseline for future UAV and terrestrial based photogrammetric studies to benchmark against, due to the high quality of the equipment which was used.

Lucieer et al., (2014) took a similar approach in their study. Whilst they collected multi-temporal photogrammetric data with a UAV, they used a different software package, COSI-Corr (Leprince et al., 2007), in order to measure surface deformation using orthophotos derived from the data. Notably, they used a series of pre-processing algorithms to produce various combinations of greyscale imagery to use within the COSI-CORR package, but were inconclusive as to whether

these manipulations affected output products. This included the first principal component in an attempt to generate a statistically optimum result, as well as using each of the colour channels individually.

2.2.3 Coastal erosion

Photogrammetry has been successfully used to make volume estimates to calculate the rate of erosion (James et al., 2013; James and Robson, 2012). James et al. (2013) demonstrated the potential for this technology by applying it to two time series collected of coastal regions, one in Sunderland, UK, and another in Uboka Cove in Croatia. These surveys were carried out using a Canon EOS450D with a 28 mm fixed focal length lens over the course of several months, allowing for the calculation of volumetric change. Westoby et al. (2012) surveyed a sea cliff at Constitution hill, UK, and results showed a good degree of consistency between the images acquired from terrestrial photogrammetry and those from a laser scanner. Considering the size of the cliff in question, and the complex topography and vegetation on it, the results show how useful photogrammetry can be for this type of research.

2.2.4 Soil erosion

Photogrammetry has been deployed as an effective tool for local scale monitoring of the effects of agriculture, land-use change and rainfall on soil erosion (Eltner et al., 2013; Kaiser et al., 2014; Nouwakpo et al., 2014; Rieke-Zapp and Nearing, 2005). This has led to a range of studies attempting to associate uncertainties derived from photogrammetric products with erosion estimates, considering the resolution required is higher than for many other geoscientific applications at a larger scale (Hänsel et al., 2016; Prosdocimi et al., 2017).

Eltner et al., (2013) used a UAV based photogrammetric survey to create a 3D mesh of an agricultural area over which terrestrial laser scanners had been deployed to generate accuracy estimates. At flying heights of approximately 10 m, using a Sony NEX-N camera with a 16 mm fixed focal length lens resulting in Ground Sample Distance if 2.5 mm, the resulting mesh showed a mean error of within 1 cm of the reference terrestrial laser scan. This is almost exactly in line with the 1:1000 rule, that the vertical accuracy can be estimated as 1/1000th the viewing distance (James and Robson, 2012).

At a finer scale, Nouwakpo et al. (2014) used very close range terrestrial photogrammetry to build 3D products of small soil plots and analyse their micro topography using a Canon Rebel XT with a 20 mm fixed focal length lens. The camera was mounted on a mobile platform travelling on a rail above the plots at a distance of 3.1 m, though this led to a non-convergent imaging geometry, which may have led to dishing/doming errors (James and Robson, 2014). The comparison in this

instance was made between a SfM workflow and traditional photogrammetry (Leica Photogrammetry Suite), with SfM having higher errors in terms of GCP residuals after processing compared to traditional photogrammetry. When each product was mapped onto one another using the ICP algorithm (Besl and McKay, 1992), they showed little difference between one another, though they note the two are deformed with respect to one another, suggesting some dishing/doming had occurred.

As with landslide monitoring, soil erosion measurements have seen increasing collection of time series data at different temporal resolution for the analysis of processes which shape soils (Eltner et al., 2015; Kaiser, 2016). Whilst Eltner et al. (2015) analyse changes in agricultural fields over the course of 10 months, Kaiser (2016) collected data of an evolving landslide every 10 minutes.

2.2.5 Structural geology

While geomorphologists are interested in photogrammetry as a means to map and monitor both vegetation and bedforms, geologists have successfully applied the same methods to generate 3D data of numerous geological phenomena (Bemis et al., 2014; Cracknell and Reading, 2014; Johnson et al., 2014; Tavani et al., 2016; Vasuki et al., 2014).

Bemis et al., (2014) argue that photo-based surveys can replace aerial LiDAR scanning where coverage is not available for neotectonic analysis. Within their study, they created a 3D model of an offset channel (Zielke et al., 2010) using just 56 photos captured using a GoPro camera on a pole. They go on to demonstrate the use of photogrammetrically-derived orthophotos for fault line mapping.

Tavani et al. (2016) utilise the technology for outcrop mapping in 3D in order to generate and analyse cross-sections. This allowed the mapping and digitization of bedding surfaces, joints and fault segments along the face of the outcrop, allowing their characterization. From this, the evolution of the faults was estimated.

Vasuki et al., (2014) flew a UAV with a mounted camera (Canon 550D) in order to survey Piccaninny Point, a coastal cliff in Tasmania, Australia. After image acquisition, they describe a novel photogrammetric supervised learning technique, where users guide the generation of fault detection and mapping within the images, using edge detection algorithms (Canny, 1986).

2.2.6 Glacial mapping

Glacial mapping presents some of the most challenging environments for optical equipment to operate in, for example debris covered glaciers, where both very bright and very dark areas may exist within the same scene. Cameras with a high dynamic range are desirable in these situations,

though for glaciers without debris photogrammetry has been shown to produce high quality 3D products (Ely et al., 2016; Immerzeel et al., 2014; Ryan et al., 2015; Whitehead et al., 2013).

Ryan et al., (2015) deployed a UAV with a Panasonic Lumix DMC-LX5 camera using a variable focal length lens set to 5.1 mm, flying at an altitude of 500 m to survey Store Glacier in West Greenland. Supplementary flights were executed at a lower height of 250 m to ensure full coverage of the glacier. The DEMs produced were sampled at a rate of ~50 cm to reduce the computational processing required.

Gómez-Gutiérrez et al. (2015) used High Dynamic Range pre-processing to investigate whether artificially increasing the dynamic range of each input image into a photogrammetric workflow could be used to enhance SfM products in a debris-covered glacial survey which required a high dynamic range. They found no improvement in the RMSE obtained when comparing the HDR and standard dynamic range images against one another when the reconstructions were compared with a reference TLS survey. This is potentially due to the fact that the camera used (Canon EOS 5D) had sufficient dynamic range to cover the scene but the authors do not report ISO so the expected dynamic range can't be estimated. Another potential explanation is the three images acquired were subject to sufficient noise that the HDR stack introduced a smoothing effect across each pixel, which potentially degraded images when compared with the exposure covering the middle of the scene's dynamic range.

They found that in two out of the three surveys conducted over the course of four years, the HDR images produced residual errors of approximately double that of the conventional images (6.8 mm vs 3.4 mm) in terms of marker error. The authors note no improvement in visual dynamic range improved this metric across the site for each of the three surveys, though do not report or comment on tie point accuracy or offer insight into why this was the case.

2.2.7 Fluvial mapping

SfM has been extensively applied to fluvial environments and for analysis of watersheds and braided rivers. For example, researchers have applied the technology for volumetric analyses of grain size at various stages along riverbeds, both above and below the waterline (Woodget, 2016). Outside of grain size analysis, SfM has also been used to measure surface processes.

Javernick et al., (2014) completed a comprehensive study of a braided river in New Zealand. They used a 10 megapixel (MP) Canon camera (the exact model was unreported) to acquire 147 aerial

images from an altitude of 600 m in order to generate and test the accuracy of a SfM derived topographic model. With 60% sidelap and 75% forelap, they reported a decimetre-scale accuracy using a SfM workflow.

Fonstad, 2014 undertook a similar study on a riverbed in the United States, using a compact camera (Canon A480) on a blimp aerial platform. 304 photographs were collected from a mean height of 40 m and used within a SfM software package in order to produce the resulting DEM, which showed accuracies with a mean difference between the photogrammetric DEM and reference TLS data of 60 cm. This is far larger an error than one would expect in normal operating conditions (4 cm might be expected using the 1:1000 rule), though they reported very windy conditions which limited the utility of the images due to blur.

2.3 Review of photogrammetric workflow

2.3.1 Camera design considerations

Camera settings, which will frequently referred to in image analysis, are the basic raw inputs to the photogrammetric workflow. In this section camera design considerations and the effect of varying each setting which users have control over will be discussed. The effect varying each of these would have on image and pixel quality will be outlined in each case.

2.3.1.1 The exposure triangle

Three settings are responsible for the degree to which a camera's sensor, located within the body of the camera, receives and processes electromagnetic radiation to which it is exposed; namely shutter speed, aperture and ISO. These can be neatly represented in an 'Exposure triangle', which shows the relationship between each of these settings (Figure 2.2).

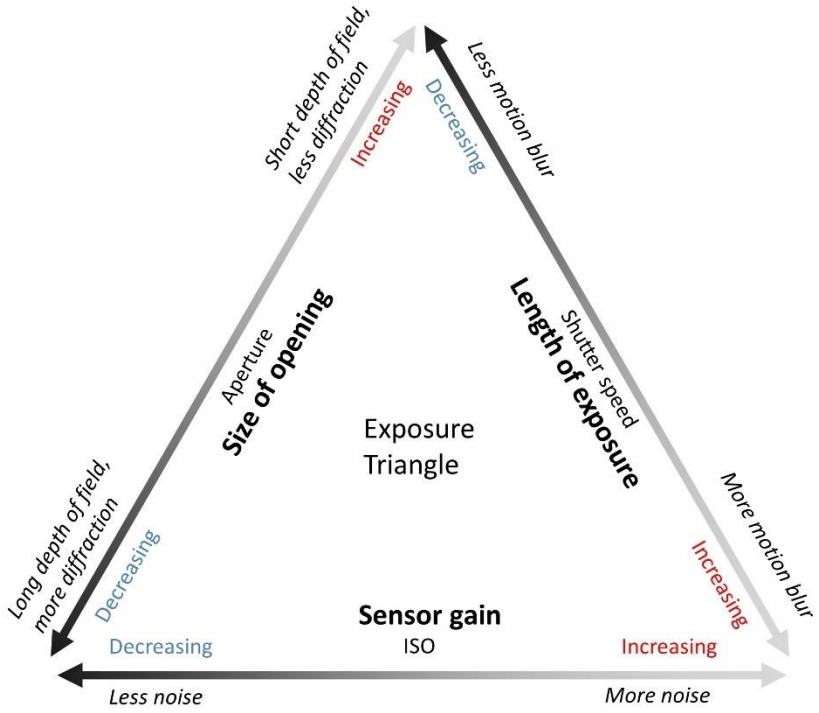

Figure 2.2. The influence of camera exposure settings on image quality is summarised by the exposure triangle. The lightness of the arrows generally reflect how the parameter will effect expose, with lighter areas representing more exposure. Image from O'Connor et al., (2017).

Aperture represents the size of the light-limiting stop located around the centre of the lens, on the same plane as the principal point. Aperture is often presented as an 'f-number', namely the ratio of the focal length of a lens to the diameter of the light-limiting gap. For example, the setting f/8 represents a gap 1/8th the size of the focal length of the lens. Increasing the aperture (decreasing the f/number) will increase the amount of light incident on the sensor, though will decrease the Depth of field (DoF). The DoF is the range in front of the camera lens within which objects are perceptibly in focus.

Thus, in many situations it is intuitive to decrease the aperture (increase the f-number) to increase the DoF. However, this can have a secondary effect of introducing diffraction effects. Diffraction is a phenomenon whereby waves passing through an aperture will spread out in a wavefront proportional to the inverse of that aperture. By increasing the f-number (N), the diameter of the Airy disc (d) also increases (Figure 2.3, derived from equation 1), thus spreading the radiation reflected from a point in object space over a larger area of the sensor's surface. At

high f-numbers high frequency edge information is lost. The wavelength of light (λ) influences the level of diffraction. Longer wavelengths passing through an aperture suffer from more diffraction due to increased interference as described in Young's single slit experiments (Tipler, 2004). Red light has a longer wavelength than blue light, and thus is diffracted more.

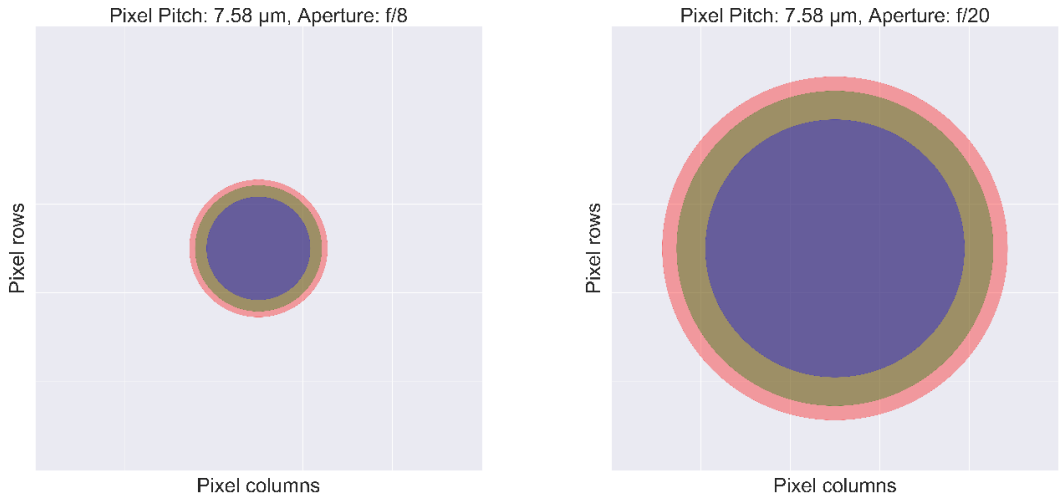

Figure 2.3. The Airy disc is the term denoting the spread of light from a lens' principal point onto a sensor. Increasing the f stop will increase the size of the airy disc, causing diffraction to occur. In the context of red, green and blue light, red light is more prone to diffraction effects than green or blue.

$$d = 1.22\lambda \mathrm{N} \quad [1]$$

Shutter speed relates to the length of time the sensor is exposed to incoming radiation. This is controlled within the camera body by a physical barrier which is inserted in front of the sensor preventing light reaching the sensor. Increasing shutter speed will increase the amount of radiation incident on the sensor, though it can also introduce motion blur which can degrade the performance of images within a photogrammetric block (Sieberth, 2014). This is particularly common in UAV photogrammetry, where wind and UAV motion can have a large effect on how blurry each image is.

In static scenes where the camera is mounted on a tripod, shutter speed can be varied more freely. Typically within the geosciences, either the platform is moving (such as on a UAV) or objects within the scene are moving (eg. wind blowing leaves), thus shutter speed is usually set to faster speeds (<1/500 s).

To distinguish colour within an image, most cameras use a colour filter array (CFA), such as a Bayer array (Figure 2.4a) which is overlain on a cameras sensor. Cameras interpolate data from pixels filtering the same wavelength of light to reconstruct a colour image from these data. For Bayer arrays, green light is sampled at twice the rate as either red or blue light. Alternatively, 'direct imaging sensors' can be used to distinguish colour by the penetrative depth of incoming radiation (Figure 2.4b)

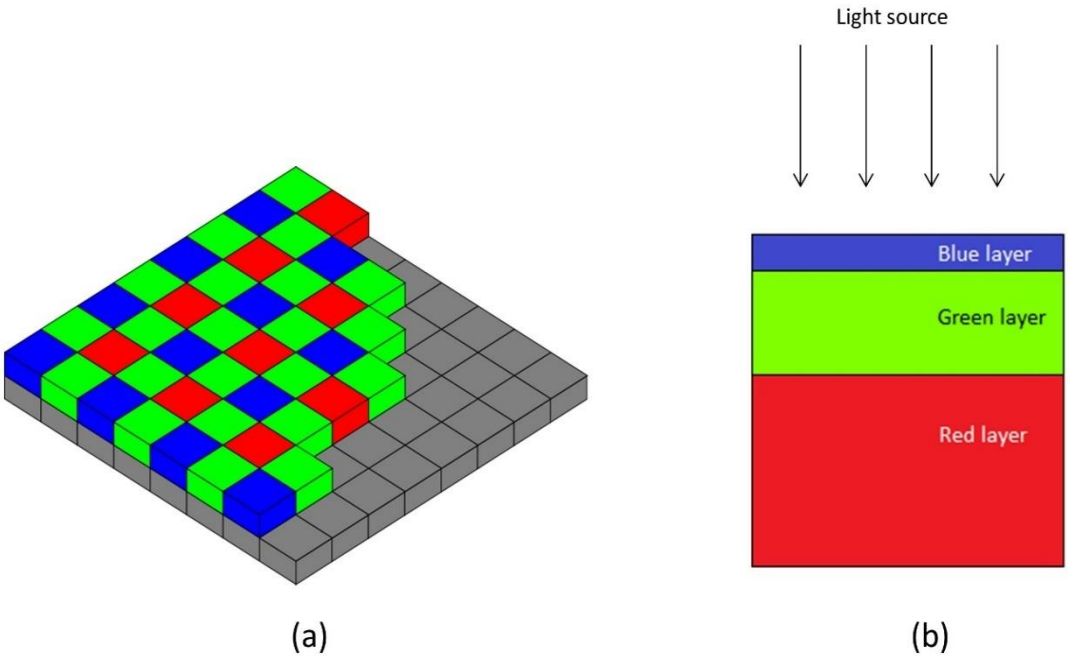

Figure 2.4 Bayer array pattern overlain on a camera sensor (a), with green pixels are sampled at twice the rate of either blue or red. Direct imaging sensors (b) distinguish light depending on the penetrative depth of the radiation. Image from O'Connor et al. (2017).

ISO refers to the International Organisation for Standardisation (International Organization for Standardization, 2006), which is responsible for standardisation across a broad range of applications. ISO, in the instance of camera settings, is a legacy term which refers to the sensitivity of analogue film, which can is broadly comparable to the gain applied to the digital camera's sensor on the raw exposure. Thus, increasing the level of gain on the sensor will increase the apparent brightness of the image, though the level of noise will also increase.

Radiation is subject to random fluctuations in the arrival rate of photons to the sensor which manifests itself as noise. Whilst dark current, the base level of noise due to variations within the camera's electronics, is measurable, the principal noise source within photographic images is shot noise. In instances where low numbers of photons are arriving (due to the scene being dark) a slow shutter speed will be required in order to ensure enough photons are incident on the sensor.

The alternative is to increase the ISO, which will apply a gain to the signal, but equally amplify the noise within the image.

Levels of noise are predictable with standardised metrics for measurement. These are usually given as a Signal-to-noise ratio (SNR) in decibels. For most consumer cameras used within reported studies, users should be able to estimate the expected SNR. For example, Burns et al. (2015) report a shutter speed of 1/180 s, aperture of f/8 and ISO of 800 for their study of SfM photogrammetry quantifying the ecological characteristics of coral reefs. Using a Pentax K5 camera, an SNR of 32 dB can be estimated, assuming the images are well exposed. One of the goals of this study is to try and objectify this information so it can be integrated into image quality analysis for intercomparison between studies in the future. Investigating this gap of knowledge intends, with the other focuses of this research, to give geoscientific researchers insight into how these decisions might impact the quality of the results they generate.

SNR is defined by the ratio of the intensity of the received signal vs. the noise inherent in the signal and that generated by the electronics of the recording device, put simply:

$$SNR = \frac{P_{signal}}{P_{noise}}$$

$$SNR_{db} = 10\log(\frac{Psignal}{Pnoise}) \quad [2]$$

$$P_{noise} = P_{signal}/10^{SNR_{db}/10}$$

Because of the very wide range of metric values noise signals can have, this is often described in logarithmic units, known as a decibels, a standard metric for use within signal processing. To add noise systematically to an input signal, such as an image, an expected noise profile based on a given signal-to-noise ratio can be produced. This noise can be added to the original image signal to degrade its quality, which is frequently done when testing the robustness of algorithms to perturbations (Rublee, 2011).

2.3.2 Camera parameters

2.3.2.1 Sensor size

Sensor size refers to the size of the sensor located within the camera body. The sensor is lined with photodiodes (pixels) which are light sensitive cells which carry charge when exposed to incoming light. The sensor size (or width, w) and focal length (f) are used to calculate the Field of View (FoV), the angle of view up to which the camera can image with regard to its nodal point,

the centre of the camera's lens. The image footprint, the area of ground the image covers, if viewing from nadir, can then be calculated, along with the Ground Sample Distance (GSD) using:

$$FoV = \frac{f}{w}$$

$$GSD = H * \frac{S_{det}}{f} \quad [3]$$

$$Image\ footprint = GSD * Number\ of\ pixels$$

'Pixel pitch' (S_{det}) refers to the size of each pixel on the sensor while H refers to the height above the ground. Larger pixel pitches are desirable as, in general, they will both collect greater amounts of light per pixel, which will give freedom within the exposure triangle and mitigate diffraction effects as light from object space is less likely to fall over multiple pixel.

GSD is a frequently utilised metric to describe this relationship, though is of limited use in situations where images are diffraction limited or where blur is an issue. Sensors of the same size can contain varying numbers of pixels, and thus number of pixels or sensor size given on their own do not provide enough information to calculate the diffraction limit, which gives some insight into the image quality.

2.3.2.2 Focal length

Focal length refers to the distance between the nodal point (the centre of the lens) and the sensor. Varying focal length will change the field of view in image space. For example, longer focal lengths will decrease the field of view and increase the GSD of the image. The focal plane refers to the plane which is located at the point of focus of the lens which is perpendicular to the lens.

Zoom lenses are commonly included within compact camera systems, though are generally undesirable as the moving parts are more prone to introducing tangential errors into the system, due to lens elements not lining up correctly. They generally perform less well than fixed focal length equivalents, 'prime' lenses, in terms of retaining sharpness towards the edge of image frames. Zoom lenses allow the focal length to be altered by varying the distance between elements within the lens body. Thus, prime lenses should be the default choice for practitioners of SfM photogrammetry, to avoid introducing additional sources of error to the camera systems.

Interchangeable lens cameras (ILCs) are designed to support different types of lenses with varying weight, focal length and aperture. While zoom lenses are commonly used with ILCs, both these

and prime lenses can be fitted to optimise flexibility of FOV and image quality depending on the survey design and objectives.

Extreme compact cameras, such as the GoPro, use fixed focal length lenses, though the focal length of the lens and the size of the sensor are both reduced when compared to ILC contemporaries. While appropriate for some survey designs, for instance where the camera needs to be extremely light and robust, they can be appropriate, though barrel distortions can be quite significant and ortorectification may prove challenging due to these limitations.

2.3.3 Image quality and compression

Within image acquisition, typical photogrammetric workflows involve capturing compressed images (usually in JPEG (Joint Photographic Experts Group) format, (Hamilton, 1992)), which involves a 'lossy' conversion where data are aggregated to reduce file size and increase write speed. However, due to the nature of the metric work, some authors have argued that RAW, or uncompressed image files, are the only scientifically justifiable file format (Verhoeven, 2010). Further advantages to using RAW include the ability to recover enhanced radiometric resolution (due to an increased number of bits-per-pixel; 'bit depth') and the ability to integrate other data, such as benchmarking tests which measure lens quality and SNR of the sensor for a given exposure into the analysis of results (DxOMark, 2018). RAW files can produce these same JPEG files, though the advantage is that the operator has more freedom in which image enhancement operations to perform during processing.

In addition, compression is not desirable. The original conditions the sensor was exposed to can never be reproduced, thus making experiments involving imagery in which environmental controls aren't monitored separately difficult to reproduce, such as Stumpf et al. (2015) who for some of their surveys captured only JPEG imagery.

Within the literature there are many proposed methods for estimating the levels of compression within images. Fan and De Queiroz, (2003) describe an algorithm they developed to estimate compression history of images. Considering the trend within photogrammetry to provide limited image metadata with studies (See appendix O'Connor et al., (2017)), algorithms such as this could present a simple way of integrating image compression metrics into geoscientific image reporting. An accurate mapping of the level of image compression to photogrammetric accuracy would present an important step in integrating image data from consumer cameras. Whilst photogrammetric accuracy can be estimated using GCP checkpoints, image quality between individual images can vary in a given study, thus these checkpoints may be sampled in images which aren't representative of the overall image quality for a given survey. Also, with the advent

of 'direct georeferencing', georeferencing which only take positional information from an image source to georeferenced photogrammetric products, thus these checkpoints aren't necessarily present.

An objective metric for image quality remains elusive, as any estimates based on either pixel or histogram values need to be sensitive to the scene which is being imaged. 'No Reference Image Quality Assessment' is a field within the computer vision community and many authors have discussed techniques which can be adopted to generate estimates (Mittal et al., 2012; Wang et al., 2002). Mittal et al., (2012) present a technique which uses the mean-subtracted normalized contrast algorithm to produce an image on which they search local luminance values. Whilst sophisticated, these algorithms have not been developed to the point where they are useful within the geosciences as out-of-the-box software does not exist.

Research reports experiments where degrading image quality, rather than compression, has shown to have a notable effect on the functioning of SfM workflows (Sieberth et al., 2013; Sieberth et al., 2014). researchers suggest processes for both detection and correction of blur, focusing errors and noise within images, these have not reached a stage where they can be easily integrated into reporting of results within the geosciences.

2.3.4 Image processing

Image processing is a signal processing method used to enhance the products of an image by extracting useful information from it. These can be for aesthetic reasons, such as modifying the colour-balance of an image to make it more intelligible for the human eye, or metric reasons, such as extracting out certain land use types from a satellite image. Within this section, image processing techniques pertinent to this work will be outlined.

2.3.4.1 Histogram manipulation/Increased bit depth

Contrast enhancement is a general term given to a range of image processing techniques for altering contrast in an image or parts of an image (Maini and Aggarwal, 2010). Typical operations include histogram stretching, histogram equalization and locally-applied variations of both. For underexposed regions in an image, shadow recovery can be used for increasing contrast within these regions (Guo et al., 2011), as contrast-stretching will have limits as noise can often dominate these regions. The use of pre-processing within SfM contexts remains understudied, with little literature existing detailing the effects these processes may have on image registration, the effects of the derived feature sets on bundle adjustment and the accuracy of the resulting models after dense cloud generation. Within some open source applications (for example VisualSfM), only 8-bits per channel JPEG images are capable of being processed.

2.3.4.2 High Dynamic Range Imaging

High dynamic range imaging (HDRI) is a technique used to artificially increase the radiometric range of a scene, and involves acquiring several images from a static position while varying the exposure (Debevec and Malik, 1997). Images with radiometric differences have been studied previously in the context of dense matching (Hirschmüller and Scharstein, 2009) but the use of HDRI in terrestrial photogrammetry surveys is not well documented.

2.3.4.3 Greyscaling

An additional step to consider in contrast enhancement is the conversion of three-band colour images (Red, Green, Blue; RGB) into single-band greyscale images. Within the SfM context, feature detectors used in sparse matching (the initial alignment of unordered images) and dense matching (refinement of a dense point cloud based on image positions derived from the sparse matching algorithm) operate on single-channel images rather than colour images (Hirschmüller, 2008; Lourakis and Argyros, 2009).

The default RGB to grey conversion which is used in many image processing/photogrammetric suites (MatLAB, VisualSfM) is based on the rec601 luma (ITU, 1995), a recommendation made by the International Telecommunication Union and an industry standard for conversion of colour images to YCbCr encoding, where Y represents the image intensity as recommended by the ITU and consists of a linear combination of colour bands with set coefficients (equation 4).

$$Greyscale\ image = R * \ 0.299 + G * 0.587 + B * 0.114 \quad [4]$$

This intensity was optimized for image display using cathode ray tube monitors and remains an adopted standard. Outside of the commonly used RGB colourspaces, different representations of colour images are presented in the literature, in particular for classification techniques and image registration. Many other colourspace transformations contain one image band with an intensity channel, which stores the greyscale information, with two other bands containing information regarding the colour content. Hue-Saturation-Intensity (HSI) and Lightness, a, b (Lab, a and b refer to coordinates representing the colour at a given point) colourspaces are examples. HSI is the colourspace of choice for many stereo vision applications, such as those presented in Fleischmann et al., (2016). It represents colour as a combination of Hue, on a 360 ° wheel denoting wavelength, Saturation, the level to which that wavelength is represented relative to blackness/white light and Intensity, the brightness of the source. They note that the saturation channel is particularly useful in image registration, as described in Paton et al. (2015), for their forest track identification application.

LAB colourspace is also used within the same study for feature generation. This colourspace uses a lightness channel to represent strength of a light source, and two channels, denoted as a and b, to encode the colour components. LAB colourspace has been used in Verhoeven et al. (2015) in combination with Principal Components Analysis (PCA) in order to enhance output products for stereo datasets of architectural scenes, though only for bundle adjustment (not in dense clouds) and only using image residuals as an objective metric.

For specific tasks, much research has been undertaken on optimising this conversion using a variety of linear and non-linear operations (Benedetti et al., 2012; Grundland and Dodgson, 2007; Lu et al., 2012; Verhoeven et al., 2015). The applications of these to photogrammetric products has yet to be thoroughly tested in a geoscience context where residual errors are generally larger than in laboratory scenes (Gruen, 2012). A summary of commonly suggested greyscaling algorithms was included in a study undertaken by Kanan and Cottrell (2012) and their impact on classification (Table 2.1).

Table 2.1 Table of operations common for conversion of RGB images to single channel images

Input set	Method	Notes
Band ratio/index	$\frac{B_1}{B_2}$, $\frac{B_1 - B_2}{B_1 + B_2}$	Red/Green Ratio Index for example would target vegetation
PC1	Coordinate rotation to find axis of max variance	Could use eigenvectors/values to indicate relevance
PC2	Coordinate rotation to find axis of second most variance	
Downsampled	Resolution is reduced post-acquisition	Could be interesting to see impact on processing times
Intensity	1/3 (R*G*B)	Simple average
Luminance	0.2126 R + 0.7152 G + 0.0722 B	Generally, for grayscaling on HD monitors
Luma	0.299 R + 0.587 G + 0.114 B	Generally, for grayscaling on non-HD monitors
Luster	½(max(R,G,B)+min(R,G,B))	Less sensitive than changes in brightness than value
Value	max(R,G,B)	Absolute brightness information

2.3.5 Image registration

Image registration is the process of mapping one or more images onto one another and is a requisite step in the SfM process before bundle adjustment can take place. Historically, image registration involved the manual mapping of high contrast features within overlapping images which could then be used to perform a transformation on one image (the 'slave') to the reference image (the 'master'). Before the application of consumer camera technology in the geosciences, images for use in metric applications contained a great deal of metadata to aid this process, such as a lens model attached to each specific camera and position and orientation parameters provided by a Global Positioning System with Inertial Measurement Unit (GPS-IMU). This allowed bundle adjustment to be performed more efficiently as the number of parameters required for the solution was reduced dramatically, with reprojection error used as a very reliable metric for the registration.

SfM's strategy for image registration more typically involves keypoint detection and localization, which requires little to no manual intervention and also no camera model or location/orientation information to perform bundle adjustment. One of the key advances that allowed this to become possible on large image sets was the advent of Difference-of-Gaussian (DoG) keypoint detectors, which enable scale invariant features to be detected within each image (Bay et al., 2008; Lowe, 1999; Lowe, 2004; Rublee et al., 2011). These allow efficient image registration within blocks where scale is not consistent, but also allow for significant redundancy in points generated, which ensures that any potential mappings have explanatory error statistics associated with them, such as image residuals (Figure 2.1). Algorithms for the accurate selection of points to include within registration after descriptor matching have been discussed extensively in the literature, with RANSAC being a popular example (Fischler and Bolles, 1981).

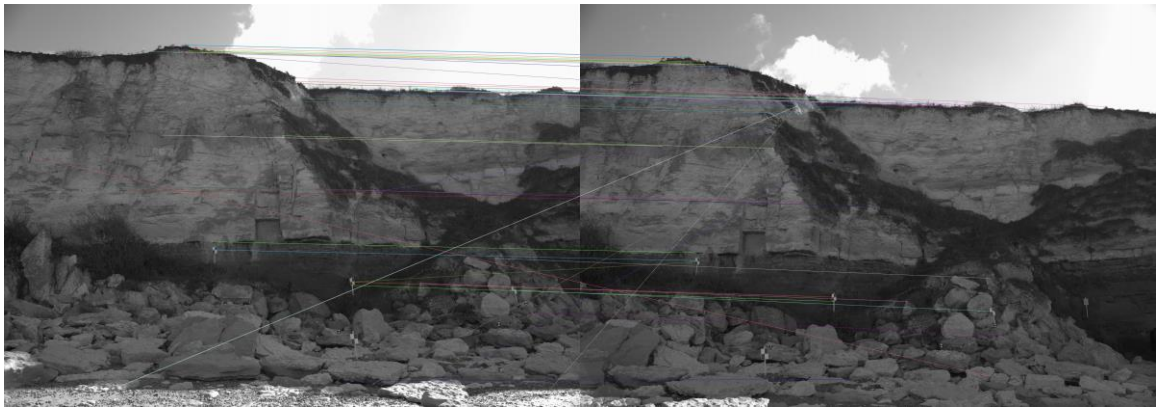

Figure 2.5 Descriptors from two images showing potential matches

One of the noteworthy aspects of keypoints which are detected for image registration using DoG feature detectors, is that features which take into account local contrast gradients are often detected at or near the scale of the minimum resolution within an image. Thus, intuitively, the contrast around a specific position would desirably show a consistent gradient throughout an image set. On flat, featureless surfaces, keypoint detection is often limited, and while certain techniques have been developed to enhance contrast and to aid the process, the literature is inconclusive as to whether there is a significant effect (Hirschmüller and Scharstein, 2009; Kanan and Cottrell, 2012). More recently neural networks have been adapted for work within image registration tasks, and represent the new cutting edge (Zbontar and LeCun, 2016).

Due to the large number of parameters required for a solution to the initial image matching and subsequently to camera positions, this process can take a great deal of time for large numbers of images. One commonly used strategy for producing solutions without the requirement for excessive amounts of computer memory and computation is to add cameras iteratively to the photogrammetric block, a strategy which has bee adopted in some commercial software (Agisoft, 2018). In an initial step, images are quickly pre-aligned at a reduced resolution to estimate potential overlaps. A seed pair with high redundancy and accuracy in this pre-alignment step is selected, and relative camera positions estimated for them. A third camera is added based on pre-alignment and fitted using the prior information calculated from the alignment of the initial pair. In practice, groups of images are added at once, and parameters in the software allow for accuracies to be selected for this process.

Bundle adjustment involves solving a system of equations which calculate relative camera positions, as well as lens models (Table 2.1) for each camera in a block. Generally, an objective function representing the residual error is minimized. In most cases within SfM software packages, this is the reprojection error, the residual error when each feature detected within in each image is reprojected back onto that image. Bundle adjustment is a non-linear process, so this objective function must be solved iteratively (Snavely et al., 2008; Triggs et al., 1999). As a result, this process leads to non-equifinality of results, thus exact results are, by definition, not reproducible exactly when using these workflows.

Konolige and Garage, (2010) detail some novelties in optimising the alignment of extremely large image blocks up to several thousands of images (so called sparse sparse bundle adjustment). This research targets datasets with sparse ‘secondary structure’; where images don’t overlap with many other images within the given dataset. While this is not yet implemented in a practical

manner for use within geoscientific research, it represents are a growing area of interest for scaling up photogrammetric work in the future.

In geoscientific studies, the accepted standard for aerial imagery is to acquire images which share an 80% overlap when viewing a surface at nadir to allow for sufficient redundancy (Smith and Vericat, 2015). Where the surface being surveyed is reasonably stable and the DoF of the scene is not significant, this standard can be used in terrestrial photogrammetric work also.

2.3.6 Camera calibration

Camera calibration involves the profiling of a camera to be used in photogrammetric work by estimating distortions within the various parts of the camera's mechanism. This is generally required to produce a solution to the bundle adjustment step, as without it images cannot be orthorectified, whereby images are reprojected such that straight lines in the real world are conserved (OSGeo, 2014).

Metric cameras have been laboratory-calibrated to solve for internal camera geometry, a so-called camera model (parameters in Table 2.2). In this Section this geometry is described, and the steps taken within the bundle adjustment stages to solve for it when a camera calibration is not provided ('self-calibration') outlined.

Table 2.2 Camera model parameters

Parameter	Description
Principal point (cx, cy)	The centre of the lens as it appears in the image, given by coordinates in pixel space.
Focal length (fx, fy)	The distance between the nodal point of the lens and the imaging sensor. These values can be different for lenses with aspherical elements, ie. where a lens is not a perfect hemisphere
Radial distortion coefficients (k1,k2)	Used for correction of radial distortion, distortions due to lens differences which are amplified towards the outer parts of an image
Tangential distortion coefficients (p1,p2,p3)	Correction for physical elements in a lens not being aligned correctly
Linear distortion coefficients (b1, b2)	Correction for differential scaling of horizontal and vertical pixel spacings and axial skew (Fraser, 2013)

The reason that straight lines on an image never appear straight is due to distortions within the camera system such as those introduced through the manufacturing of the lens. More dramatic lens distortions are obvious in cameras with extremely wide 'fish eye' lenses, where images appear distorted and curved towards the outer parts of the image, so called 'barrel' distortion. Barrel distortion generally is linearly related to the distance from the centre of the lens in a radial pattern, and therefore by creating a camera model with an estimated level of distortion built in, these latent errors can be corrected for - the process of lens correction.

Brown, (1966) describes correction of radial distortion, of which barrel distortion is one type, by fitting various coefficents to a system of quadratic equations. Simultaneously, corrections for errors in alignment of optical parts were included in the equations, meaning the whole lens model could be solved for simultaneously. This model is still widely used today and is the basis for the discussions of camera calibration.

Zhang, (2000) first described a convenient method for estimating these distortion parameters using a chequerboard target and a camera. The core concept lies in the knowledge the lines on the target are straight, and so by integrating information from the same camera from different positions (after image registration), we can add in more prior information, ie. that the straight lines are straight. The software developed uses a corner detector to detect the corners of each chequer of the target and from this information can estimate distortion levels where concurrent sets of points are known to be in a line. This idea is based on the principle of epipolarity, the geometry of stereo pairs, and homography, which describes the relative mapping of one camera to another using a transformation matrix. A second type of distortion, known as pincushion, leads to an optical effect where the object space appears magnified in the image space, again increasing radially out from the centre of the lens. This type of distortion is generally associated with telephoto lenses (long focal lengths) which are rarely associated with cameras used in the geosciences, though some exceptions exist (D'Amato et al., 2016). While Zhang's method for generating camera models might be advised in certain situations where accurate calibration is required for the functioning of the system, such as two cameras mounted on a stereo bar, it is impractical for ad-hoc geoscientific studies where lab-based calibration is not deemed a requirement.

Other parameters present in the lens' model solutions include the location of the principal point. This is required due to errors in both the manufacturing of camera sensors and lenses, which mean that the centre of the lens and its focal point are not always aligned with the centre of the camera's sensor. This can introduce asymmetric errors within solutions to the lens distortion

coefficients, though by estimating any offset these effects can be mitigated (such as proposed in Brown's camera model).

Tangential distortions are caused by elements within each part of the lens/camera not lining up along the axes perpendicular to the propagation of light, such that parts of the system are slightly tilted with respect to one another. This so called 'axial skew' is often slight, and commonly omitted from lens models when using consumer grade cameras to mitigate the risk of over-parametrization (Nouwakpo et al. 2014, James et al. 2017a). Lastly, some lens models allow for the correction of errors in aspect ratio.

Certain errors are not dealt with at the bundle adjustment stage but may be estimated and integrated into the camera calibration workflow before bundle adjustment. Chromatic aberration is a natural phenomenon caused by the focal points of light of different wavelengths in an optical focusing system not falling on the same plane. For example, for a sensor/lens combination which has been calibrated so that the focal point is optimised for green light, blue light will fall slightly in front of this, causing astigmatism in the blue channel. Conversely, in the red channel, the opposite effect will occur. This is notable in images with high contrast edges where 'colour banding' is evident. Sieberth et al., (2013) estimate that this can cause an estimated 1/20th pixel error in registration, mitigation therefore is desirable in large photogrammetric blocks.

Moiré is another optical effect which is not explicitly dealt with in the camera calibration phase, though is notable within single images. It is caused by the sensor sampling an image which resides at or near the sampling frequency of the sensor itself, known as the Nyquist limit. When sampling at the Nyquist limit, pixels containing high frequency information which is slightly out of *phase* with the camera sensor will produce convoluted pixels which are blocks of a wavy pattern. This is also known as 'aliasing'. While moiré is typically seen more in man-made scenes and rarely in nature, it might become a consideration in specific geoscience situations when imaging a highly regular pattern, such as extremely ordered crops in a field.

Typically, SfM software packages fit lens models as part of the solution without prior information of real world geometries, or of the camera lens model. This is possible due to the amount of redundancy typically present within photogrammetric blocks. Multiangular information adds enough to make the simultaneous solutions possible, though can sometimes impact them if care isn't taken in planning the camera network correctly (such as dishing/doming, James and Robson, (2014)). Camera calibration is more accurate in situations where a 'strong' network geometry is adopted, where the optical axis of each camera is not always orthogonal to the surface normal (Wackrow and Chandler, 2011). Other optical effects at the imaging stage, such as chromatic

aberration can potentially be mitigated during a pre-processing stage (Cronk et al., 2006), for example using chromatic aberration distortion coefficients specific to each camera in the image processing software. Mitigating moiré is impractical, pre-planning to avoid the problem is recommended. Fujifilm have developed a specific sensor aimed at minimising moiré effects by disrupting how regular each colour is sampled on the array and thus avoiding issues with the Nyquist sampling limit at the cost of a lower sampling rate of Red and Blue information (FujiFilm, 2018).

Other factors exist in calibrating cameras which are largely ignored by the geoscientific community, but potentially have impacts on photogrammetric products. These include the noise level associated with the sensor (Hasinoff, 2014), and the effect of exposure on results (though is included in computer vision research, such as Middlebury, (2015)). Typically, the information required to estimate these parameters (illumination level, camera ISO, shutter speed and aperture) are not included in the results of geoscientific work (see appendix O'Connor et al., (2017)). These can both be pre-calibrated to give an objective reference, though this is often passed over due to its inconvenience and requirement of technical equipment (such as an integrating sphere).

The noise levels of various cameras are measured objectively by various benchmarking organisations (such as DxOMark), which allows for speculation as to the effect of sensor quality on photogrammetric outputs. Debevec and Malik, (1997) described a method for spectrally calibrating an image such that a radiance map, displaying the scene using SI units (W / m^2). These considerations remain outside the scope of the community as it stands and will be considered within the forthcoming tests.

2.3.7 Network design

Network design describes the position and orientation of cameras within a photogrammetric block, and has been well studied within the photogrammetric literature in both industrial and geoscientific contexts. Mason, (1995) discusses network design and presents four fundamental considerations (Grafarend, 1974). These are firstly datum definition; how an operator will define the datum for use for a particular survey (Zero-order design, ZoD). Secondly, where and how many cameras to place within the network (First-order design, FoD). This is undoubtedly the most difficult consideration and is often undertaken using simulations – a network is simulated, and a bundle adjustment run to estimate precision. Thirdly, observations are weighted if required. Lastly network densification if required. Fraser, (1984) outlines these considerations which show the effectiveness of convergent network design, with precision estimates from convergent

networks returning consistently lower values than 'traditional' aerial surveys with nadir pointing images.

Rules-of-thumb have arisen from recent research, with separate strategies for both aerial and terrestrial data acquisition to consider. SfM literature detail some of the most obvious forms of systematic errors as a result of incorrect camera calibration, as well as outlining intelligent network design, in particular with regards to FoD optimisation (James and Robson, 2014; Wackrow and Chandler, 2011; Wu, 2014).

Many contributions attempt to improve camera placement in network design, such as Olague, (2001), who attempts to optimise the FoD solutions presented in Fraser, (1984) by using different methods for solving for precision, in this case genetic algorithms. Ahmadabadian et al., (2013) suggested a novel algorithm to aid in FoD decision making, based on Hanel et al., (2012), who suggested an iterative algorithm for adding cameras to a scene in order to capture complex geometries. Whilst a rough model is required to initialize the algorithm, the potential offset of additional images is in many situations worth the effort, for instance optimal camera localisation for 3D time series generation, as seen in Eltner et al., (2017). Within a geoscientific context, camera placement will often be met with these constraints, making the simulation approach more difficult to implement practically.

Hanel's paper decomposes the practical challenges associated with optimal structural capture using a single camera into three sections: range-related, factors regarding the focal length, sensor and distance to the object of interest, visibility-related, incidence angles, occlusions and camera field of view and accessibility-related, where images can practically be acquired from.

2.3.7.1 Range-related issues

Whilst the idea of practical limits on focal length were introduced in an earlier section, James and Robson (2012) explicitly note the difficulty of fitting a correct camera model when using longer effective focal length lenses (>35 mm). Within the paper, the concept of the maximum achievable coordinate precision is introduced through an equation from (Fraser, 1996),

$$\sigma_c = \frac{qD}{\sqrt{k}d}\sigma_i \quad [5]$$

where *q* is a factor that represents the strength of a photogrammetric network geometry, D is the mean distance from the camera to a target, *σi* is the precision of image measurements and *d* is the principal distance of the camera. They use nominal values of k = 3, q = 1 and an effective focal length of 32 mm to estimate the practical limits of coordinate precision as being roughly 1:10000

in terms of estimated precision to viewing distance. Thus, for a surface 50 m away, a maximum achievable coordinate precision is estimated at 5 mm.

With these constraints, it becomes clear that one should try and use the longest focal lengths practical for a given objective (increasing d in the equation will make coordinates more precise), as discussed in O'Connor et al., (2017), though this must be balanced with distortions introduced as a result of these long focal lengths.

2.3.7.2 Visibility related issues

A 'strong' network design is desirable to mitigate potential systematic error, though it is not always possible to acquire, for example where certain locations are inaccessible. Occlusions include anything which might obstruct the camera's view of the surface of interest, and are quite often unavoidable in a geoscientific context, for example when a bare Earth Digital Terrain Model (a topographic representation of the surface when stripped of vegetation) is being acquired from an aerial platform and vegetation is occluding the surface. Lastly, in areas where access is limited, camera operators may not have adequate equipment, such as short focal length lenses when up close to an object, to adequately survey an area.

2.3.7.3 Accessibility related issues

This refers to the ease of access to a site for a survey, or sites where access is restricted. In the geosciences, it is common for certain areas to be inaccessible, such as sea cliff faces. UAVs are used to overcome accessibility constraints (Genchi et al., 2015).

2.4 Summary and knowledge gap

This literature review has introduced the most common uses of photogrammetry within the geosciences and give an overview of the principles of SfM photogrammetry and the workflow involved. Issues regarding image quality and image pre-processing were introduced, to contextualize these issues which are discussed within the remainder of this text. Lastly, network design was discussed, from which key points are drawn in the research design of this contribution.

The most important consideration introduced, in the context of this research, is the gap of knowledge with regards image quality issues and their impact on image registration, bundle adjustment and dense cloud generation processed within SfM. Owing to the highly variable quality of images/cameras used in SfM research in the geosciences, the focus of the research portions of this text are to highlight how this affects the products which geoscientists generate from this workflow.

Chapter 3: Methods and research design

3.1 Introduction

This chapter will be split into three parts. Firstly, several proposed research questions will be addressed. Following from this, the technical requirements to answer each question will be discussed and an appropriate research design presented. Secondly, the research sites will be described, with reference to the technical requirements presented in the first section. Additional datasets used within this work will also be presented within this section. Lastly, the data handling and processing strategies will be described in detail, with each quantitative metric presented justified.

3.1.1 Research questions

3.1.1.1 Does a decline in image quality reduce photogrammetric accuracy?

For researching effects of image quality within photogrammetric processes, one would ideally collect 'high quality' and 'low quality' image data in the field for direct comparison within photogrammetric processing. One difficulty with this step was ensuring static camera positions between sets of images. This is important to ensure that differences between the derived products (image matches, point clouds) are independent of differences due to environmental conditions, such as a change in illumination or wind causing motion blur.

Within this research project a standard reference set of images were collected and image quality was degraded by manipulating the RAW images within the processing steps proceeding photogrammetric processing. These degradations include the addition of JPEG compression and addition of noise (typical of high ISO values). This strategy simulates low image quality due to poor data acquisition strategies, and provides insight into the potential pitfalls of using poor quality data within photogrammetric processes.

3.1.1.2 Does the greyscaling algorithm used affect the results of sparse/dense point cloud generation?

Most consumer cameras have a colour filter array laid over the sensor's surface to sample red, green and blue light. SfM algorithms generally use a single grayscale channel to perform image registration. As the green channel is sampled at twice the spatial resolution as either the red or blue, improvements in both image matching and dense cloud generation, due to the increased sampling rate would be expected. To formally test the effect of utilising just one band, subsets of grayscale images were derived from the camera's RAW imagery. Results from these image sets

were then tested against one another, as well as the original untreated image sets, to establish whether any differences were observed.

To test whether colour variation affects the image matching, a naturally colourful scene was required to ensure colour variation. For a monochrome scene, it is intuitive that using only the green channel would be best, though where high frequency information is stored within the colour bands this may not be the case. The objective of the research design is to establish if this effect is detectable.

3.1.1.3 Does the JPEG quality influence image matching and dense cloud accuracy?

JPEG processing has become a ubiquitous part of the workflow on-board consumer cameras, though the effect of this compression has not been formally tested in the literature. Considering many studies within the literature do not report the type of image data utilised, we assume them to be JPEG images. JPEG algorithms involve a discrete cosine transformation using look up tables as described in the previous chapter.

These tests require a set of high quality reference images, as with section 3.1.1.1, against which a set of systematically compressed images can be compared for accuracy and point density.

3.1.1.4 What effect does resolution have on accuracy of points?

To test whether there is a correlation between the number of image pixels and the accuracy and density of the derived point cloud, images were steadily downsampled to different pixel grid sizes, before undertaking photogrammetric processing.

3.1.1.5 Does the added bit depth of TIF files enhance accuracy in sparse/dense matching?

Bit depth, as described in Chapter 2, is the number of bits the pixel intensity data are quantized to once read off the sensor. RAW images from COTS cameras can write anywhere from 10 to 16 bits per pixel, and this experiment tests whether this influences the accuracy and density of products from photogrammetric processing.

From the high-quality reference RAW data, TIF files, capable of being handled by most photogrammetric software packages, were derived. These photogrammetric products were measured against the highest quality JPEG sets available to establish any difference within density/accuracy.

3.1.1.6 How does viewing geometry affect both accuracy and density of products?

Accuracy of photogrammetric products is closely linked to the strength of the camera network, with convergent imaging configurations recommended to help ensure fidelity of derived products. To test the effect that imaging geometries have on products of photogrammetric processing, sets

of images were systematically added to the SfM workflow with varying numbers of images and viewing angles.

3.1.2 Technical requirements

Within image acquisition and SfM photogrammetric workflows, users have a large number of parameters which they can vary depending on the equipment and software used (Table 3.1, 3.2). For some, users can have near full control (such as ISO, shutter speed and aperture), though there are several which will only be estimated prior to performing a survey (such as the exact camera positions images will be acquired from). Using error metrics at various stages of the processing chain can indicate how these parameters are affecting derived products, such as image residuals indicating accuracy of image matching, and dense cloud densities/point counts providing feedback on the performance of the dense matching algorithm.

Table 3.1. Considerations for the camera, camera settings, camera network design and data validation when designing the fieldwork part of a survey.

Parameter	Control
Sensor size	Manufacturer standard
Aperture	Function of the selected lens/User defined
Focal length	Function of the selected lens
Shutter speed	Controllable
Distance to target	Controllable
Camera positions	Can select, though not surveyed. Static camera orientation on a transect is desirable, though not required.
Reference surface model	Attainable through a TLS survey, referenced to the photogrammetric outputs by Total station surveyed Ground Control Points (GCPs). Errors will exist in both the TLS and Total station surveys due to limitations of the sensors. Occlusions may differ between the photogrammetric survey and reference model.

Table 3.2. Considerations for post-acquisition processing and validation.

Parameter	Control
TIF image generation	Can optimize based on image statistics and apply the same operation on all image sets prior to further analyses
JPEG image generation	Controlled using a standardized software package with static algorithms
Sparse bundle input parameters	Can be set to the same input parameters for each image set input
Dense cloud generation	Can be set to the same input parameters for each image set input
Camera positions	Floating variable between SfM solutions
Distance to GCPs	Error reported within software package
Distance to reference surface	Independent metric acquired from separate software package

3.1.3 Research design

The objective of this research project is to ascertain any potential improvements or degradations in SfM products associated with image quality. The goal is a set of recommendations for image acquisition and processing strategies depending on the accuracy and resolution required for a specific application. Examples include the ability to confidently select a given sensor and camera lens for use, given specific survey requirements, or use compressed JPEG imagery for a given application, knowing it may not perform as well as RAW imagery.

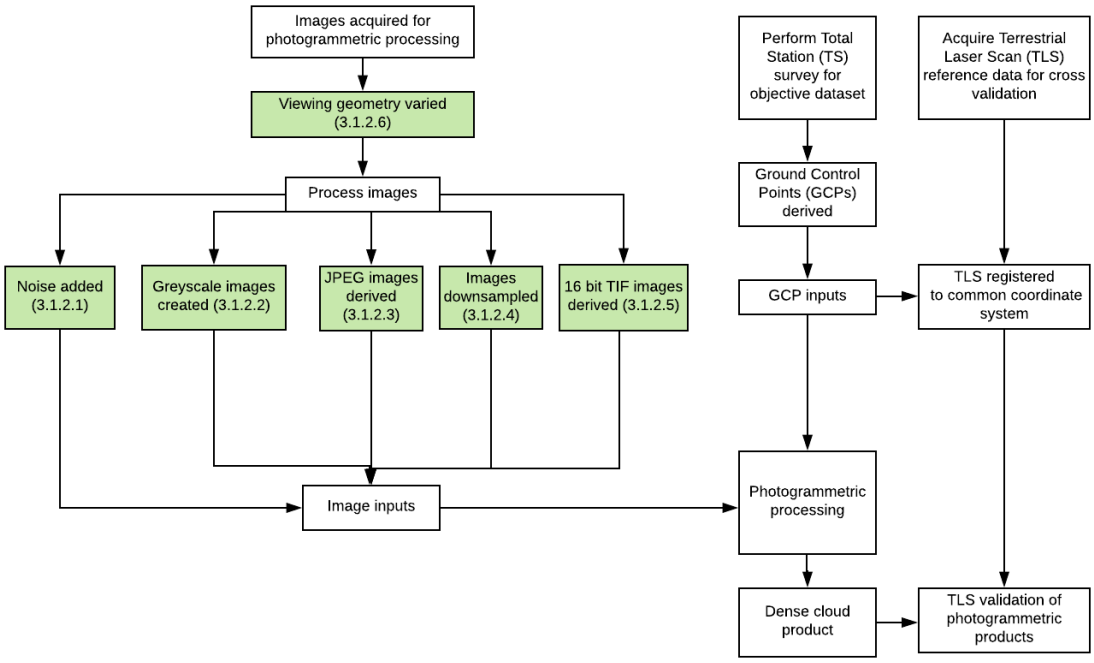

Figure 3.1 Research design. Each research question presented in 3.1.1 is addressed in the steps with green backgrounds.

3.2 Software

To answer the research questions detailed in the previous section, several software packages were required to ensure that results from the workflow were well defined, and the process could be repeated across multiple treatments with ease. The rationale behind using each package for this research is outlined within this section, with reference to the workflow described in Figure 3.1.

3.2.1 Preprocessing: ImageMagick

For image pre-processing, ImageMagick was chosen for the RAW-JPEG conversions as it is built on a software package (DC-Raw, 2018) which had been previously used in SfM research (Bettio et al., 2013; MacDonald et al., 2016). ImageMagick is an open source C++ library which provides

command line tools to run image processing applications in a scripting environment. The reason it was utilised for this research is that for each image operation considered within the research questions (RAW-JPEG, greyscale, compression, downsampling) the package has command line parameters specifically targeting each of these which could be exploited ('magick', 'fx' and 'quality'). Another consideration was the ease in which it handles data with bit depths larger than 8-bit JPEGs. For example, to perform a RAW-TIF conversion from Nikon's NEF raw image format, the command is as simple as

```
magick example_image.NEF example_image.TIF
```

which conserves the bit depth of the pixels, with automatic colour adjustment. Another example of a conversion of a RAW image file to a compressed JPEG image file at a compression level of 'quality' 50 is:

```
magick example_image.NEF -quality 50 example_image.JPG
```

3.2.2 Photogrammetric processing

The list of photogrammetric software packages available has grown significantly with the development of SfM as a topic. A list of the most common SfM packages is presented in table 3.3, with short descriptions and information given for each solution.

Table 3.3 List of SfM software solutions

SfM solution	Type	Description
Photoscan	Commercial	Agisoft PhotoScan is a SfM package which has seen frequent use across the geosciences (Javernick et al., 2014; Ouedraogo et al., 2014). It includes the ability to export all camera models, sparse point clouds, dense point clouds, DSMs and rasters, which represent a large proportion of the products which were utilised in this work.
Pix4d	Commercial	Pix4D represents one of the main commercial competitors to Agisoft PhotoScan (Strecha et al., 2008). Their product is more expensive, though they have a large client base amongst geoscientific consultation companies, owing to the focus on high quality aerotriangulation and levels of redundancy built into the bundle adjustment solutions.
Photomodeller	Commercial	Offers end-to-end SfM solution with minimal expertise required, as with Photoscan and Pix4D
Autodesk 123D Catch	Free to use	Solution for generating 3d models from unordered images and used in some geoscientific studies (Micheletti 2015). Discontinued.
VisualSfM	Open source	Built on the "Bundler" SfM package (Snavely et al., 2008) which performs image matching and bundle adjustment, and with distributions containing a dense point cloud generating algorithm (CMVS2, (Furukawa and Ponce, 2010)). While the main branch has seen development ceased, the high generality of the sparse bundle adjustment algorithm, as well as it's highly customizable parameter sets and feature trackers have meant that its use remains widespread (Mlambo et al., 2017).
MicMac	Open source	MicMac is a free, open source software which was developed by the French Forestry Institute, with a specific focus on their high quality aerotriangulation software APERO. Bolted on to this are separate modules for generating lens models, dense clouds and meshes. Used in Stumpf et al. (2015).
Theia	Open source	Theia is a library containing one click SfM solution that was designed to be easily extended to incorporate new algorithms.
Python Photogrammetry Toolbox	Open source	The python photogrammetry toolbox is a group of python scripts which automate the SfM pipeline - formerly OSM-bundler.
COLMAP	Open source	SfM solution with graphical and command line tools, as with VisualSFM.

Due to its wide use within the scientific community and this simple and extendible Python API, PhotoScan (Version 1.3.1.4030) was chosen as the SfM software package for use within this research.

3.2.3 Point cloud post-processing: CloudCompare

CloudCompare is a suite of 3D point cloud editing and processing software (CloudCompare, 2018). It contains many tools commonly found within point cloud software, such as the ability to do point selections, cropping, subsetting and cloud-to-cloud registration, amongst other operations. It is invaluable in the context of geoscientific studies, where metrics such as rate of change between point clouds acquired at different times need to be assessed. In the context of this research, its ability to generate error metrics when comparing two clouds (cloud-cloud distance using the Chamfer matching algorithm (Barrow et al., 1977)) was one of the central criteria required for a post-processing package. It also features a command line mode, which is useful for standardising operations across multiple datasets. Thus, in combination with scriptable components from ImageMagick, PhotoScan and CloudCompare, the workflow outlined in Figure 3.1 could largely be automated, which made reuse of components much simpler. Thus, the scope of the research could be widened to answer all the research questions presented in Section 3.1.1.

3.3 Site descriptions

3.3.1 Suitability considerations

Practical considerations

Ideally, the selected sites should be accessible, owing to the large amount of equipment to be transported to fulfil the goals of the research design. The feature of interest should be visible from all desired angles detailed in the survey design, with no occlusions. Given the precise measurements required ideally all equipment used would be kept static for the duration of the survey. Thus the location should be sheltered.

Research oriented considerations

One of the research questions (3.1.1.2) presented is specifically dealing with processing colour in RGB images with a geoscience focus, thus the scene would ideally include a range of colours. Coastal cliffs offer not only a diversity of colour, particularly in partially vegetated regions and areas with exposed rock strata, but are also largely well studied in areas prone to change and erosion. They offer the added advantage of allowing images to be acquired both from the direction of the surface normal and at various angles to it from a tripod mounted camera, which led to them being chosen as the geomorphological feature to be studied in this contribution.

Table 3.4. Basic site requirements for carrying out this research

Site requirement	Notes
Sheltered	Allows for equipment to be kept static for the duration of the survey
Accessible	Allows for loading/unloading of equipment, access to multiple angles for convergent viewing
Wide range of colours	Desired for testing greyscale algorithms (3.1.1.2)
Terrestrial capture	Vertical orientation of the surface

3.3.2 Pilot study 1: Lab survey

A lab survey was carried out to act as a pilot study for the fieldwork. The lab test consisted of a bed of gravel, with relatively little colour contrast, on a white background 3.2 m by 2.1 m. Nine checkerboard targets were placed in the scene to coregister the photogrammetrically derived products with those attained from a laser scanner. The checkerboard targets were large relative to the scale of the scene (full sized A5 pages) and were not held in place. The large targets were used to ensure the GCP positions could be accurately estimated in every input image, as some images were at extremely shallow angles relative to the sensor. A Canon 500D camera fitted with a variable 18-55 mm lens was used to acquire the images. The focal length was fixed at 18 mm for the duration of the survey, to maximize field of view in confined laboratory conditions. A Faro Edge hand driven laser scanner was used to generate a high quality reference surface which the photogrammetric products would be compared against (Faro, 2015). This laser scanner can attain accuracies of 24 microns, which is far below the GSD of the images from the photogrammetric survey, which averaged 260 microns. Thus, the laser scan reference is expected to be of a much higher quality than that of the photogrammetric products.

Figure 3.4 Lab experiment setup.

3.3.3 Pilot study 2: Desktop research

Aside from the research questions presented, the reproducibility of photogrammetric results needed to be established before conclusions could be drawn from the results which would be generated. To do this, a sample dataset was acquired (Westoby et al., 2012), which consisted of a set of images and a reference laser scan which acted as a reference surface. The images within the set cover the whole cliff and vary in distance, but were designed to ensure no areas were occluded on SfM execution. For these tests, the workflow designed to be run across the data collected from the fieldwork in this contribution was run multiple times across the same sets of images from the same sample dataset, in order to establish how consistent the results were, as SfM is a non-linear process which therefore suffers from non equi-finality of results.

The site, as described in the paper (Westoby el al., 2012), is Constitution Hill, a 141 m sea cliff located in Aberystwyth, Wales. It serves as a benchmark dataset on which to refine the initial investigations, as both the sites chosen in this work are sea cliffs.

The images used from the sample dataset were acquired in JPEG format only with no metadata as to the level of compression, thus were not included in image degradation experiments. These images were subsequently deemed inadequate to use as a primary data source for this research

but are adequate to serve as a pilot study for assessing the repeatability of a given set of input parameters to the SfM workflow.

Figure 3.5. A sample image from the dataset acquired by Westoby et al. (2012).

3.3.4 Field survey site 1: Hunstanton (Grid reference TF 67420 41739)

Hunstanton is a seaside town in Norfolk, England, at which a long stretch of steep coastal cliffs are found (Figure 3.6). The cliffs consist of several well-exposed layers of topmost white/grey stone known as The Ferriby formation (Grey, soft, marly chalk (BGS, 2016)), layers of red chalk known as the Hunstanton formation (Rubbly to massive chalks with marl bands; typically pink to brick-red (BGS, 2018a)), and a brownish-green sandstone known as the Carstone (medium- to coarse-grained and pebbly in part, especially at the base where it becomes a conglomerate (BGS, 2018b)). The presence of vegetation is limited, but some pioneer vegetation was present in the survey area. A 48 m horizontal stretch of the cliff was selected for survey, owing to its rich colour contrast, the presence of vegetation and complex debris.

Figure 3.6 Hunstanton panorama

3.3.5 Field survey site 2: Overstrand (Grid reference TG 25720 40619)

Overstrand is a village approximately 60 km East of Hunstanton and features many prominent coastal landslides with complex topography (Figure 3.7). Its cliffs are a Special Area of Conservation under the European Commission's Habitats Directive. The site at Overstrand consists of a 72 m section of vegetated sea cliff, characterized by frequent cliff failure and landslips. Geologically, it is made up of chalk masses at a site referred to in the literature as the 'Overstrand Hotel' (Wood, 1967). It exposes flinty chalk in the upper section with well-developed marl seams visible towards the base of the formation (DEFRA, 2018). It is of a comparably larger scale than the first site, with the cliffs 35 m above the level from which surveying took place. A central gully is a significant feature of the site, which runs down from the top of the cliff, leading to very complex topography in the middle Section of the site. The green vegetation contrasts with the deep brown mud on which it sits, and also the yellow soils seen at the top of the cliff. Viewing from ground level, some areas are occluded, such as flat areas of grass located approximately half way up the landslide. While this is not ideal (as in section 3.3.1), the site was chosen as enough other attributes fit with the objectives of the research.

Figure 3.7 Overstrand landslide panorama

3.3.6 Survey planning

Initially, low resolution 3D point clouds for each site were acquired from the UK national LiDAR inventory (UK Environment Agency, 2018) for use at the planning stage, which allowed camera specifications and requirements, as well as sampling distances and necessary equipment to be prepared prior to fieldwork execution. These data were not used past this initial equipment and design planning stage.

One requirement of the survey was that the photogrammetric targets needed to occupy at least 5 image pixels on their shortest axis (Dold, 1996; Shortis and Seager, 2014). A 35 mm lens was selected (James and Robson, 2014; O'Connor et al., 2017), to ensure targets could be surveyed at sufficient resolution. At this stage, the sampling distance from each cliff was estimated to be 30 m at Hunstanton and 50 m at Overstrand, whose extents could both be captured using a 35 mm lens on a full-frame camera, which has a vertical field of view of 37.8° and a horizontal field of view of 54.4°. A 12 MP sensor (Nikon, 2017) was selected, leading to spatial sampling rates (ie. Ground Sample Distance) for the photogrammetric survey of 7 mm for Hunstanton and 12 mm for Overstrand. Using laminated A4 targets this led to target sizes of a minimum of 15 pixels (for the targets located towards the back of the Overstrand site). A full list of equipment required for the field work was drawn up and is shown in Table 3.5.

Table 3.5. List of equipment used for fieldwork

Instrument	Used in
Nikon D700 (Nikon, 2017)	Photogrammetric survey
Nikon 35 mm lens	Photogrammetric survey
Leica P40 Terrestrial laser scanner (Leica, 2017a)	TLS survey
Tilt and turn TLS targets	TLS survey
Leica TCR 805 Total Station (Leica, 2017b)	Total station survey
Leica retro reflective target	Total station survey
Photogrammetric targets	Photogrammetric and total station surveys

3.4 Data acquisition and processing

Initially, all targets were placed in the scene, in accessible locations, and spread out to encapsulate the extent of each survey (Figure 3.4 and Figure 3.5). For the Hunstanton site, some targets were placed in front of the bulk of the landslide, but were still visible from enough images to act as ground control.

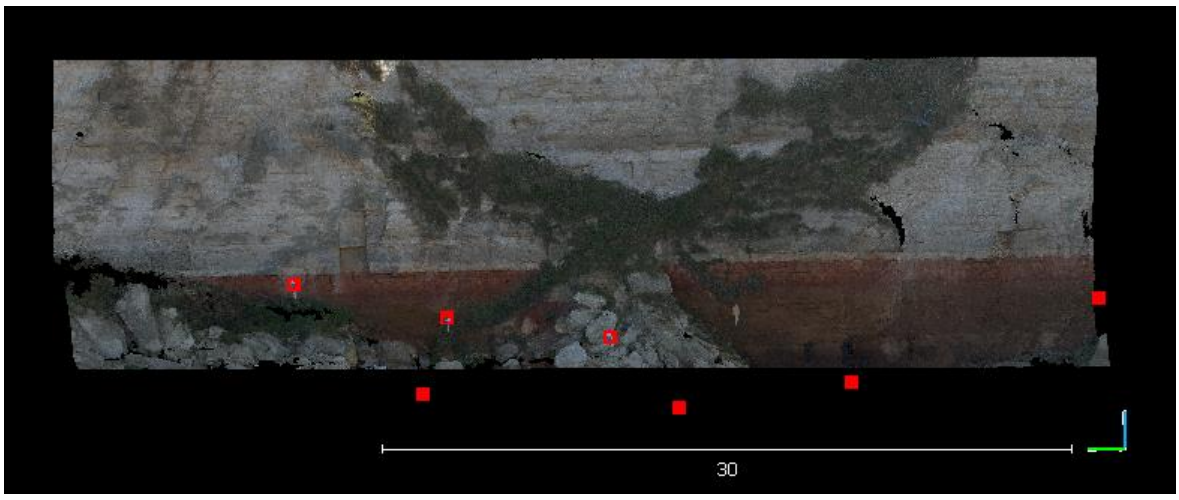

Figure 3.4 The location of the Hunstanton GCPs are shown as red points. Scale shown in metres.

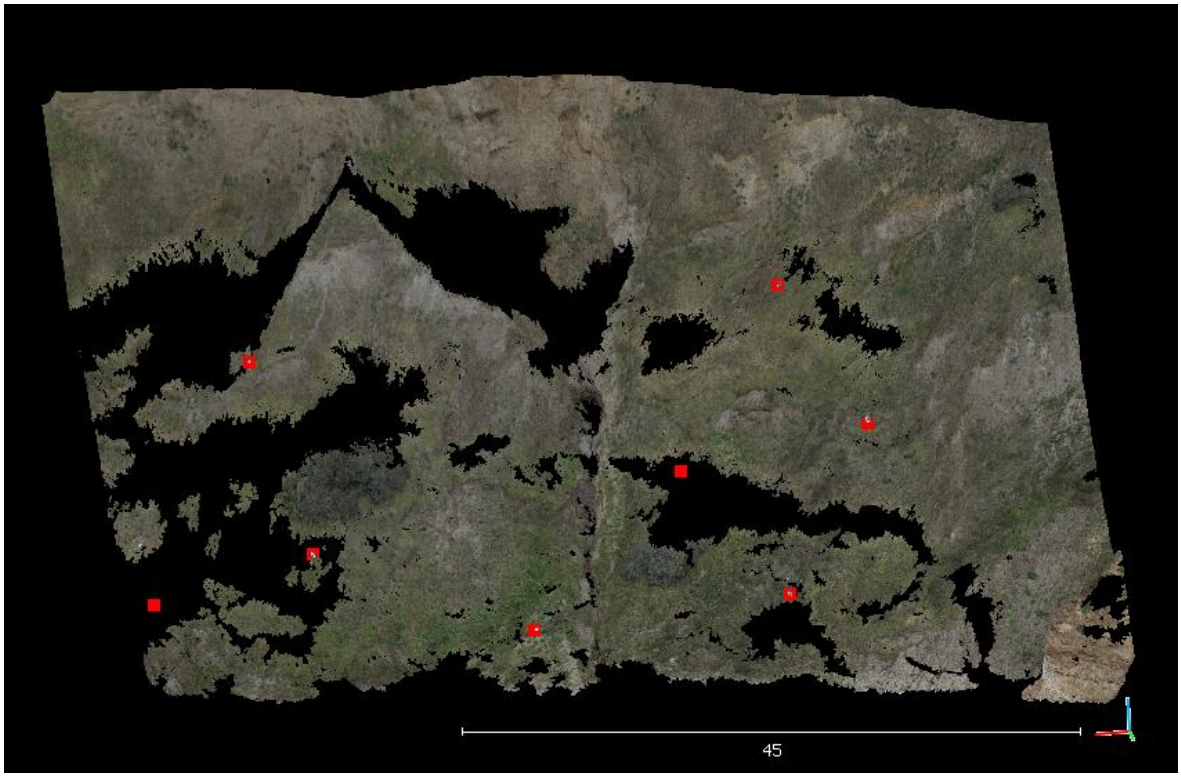

Figure 3.5 The location of the Overstrand GCPs are shown as red points. Scale shown in metres.

In total, 4 tilt-and-turn Terrestrial Laser Scanning checkerboard targets were spread out to ensure captured scans could be coregistered to one another. Following this, 8 custom made staked signs, with laminated A4 checkerboard targets glued on, were mounted in accessible positions. These targets were hammered into the ground to ensure they were static for the duration of the surveying.

3.4.1 Total station survey

The total station survey consisted of setting up a static tripod at either end of the camera transect, upon which the total station was mounted. The total station was mounted, levelled and zeroed horizontally against a retro-reflective target on the tripod at the far end of the transect. This allowed an accurate reading of the length of each transect, and ensured that measurements were consistent between each survey position.

Each target was then measured from each station three times, with the telescopic sight being reset each time to attempt to ensure fully independent readings. Readings were written down on a field notebook, and stored on the total station's internal memory for recovery later, ensuring a backup was kept. In total, 24 readings were taken from each survey position (3 for each of the 8 photogrammetric targets), these would be coregistered at a later stage to produce a final set of

GCPs for use in both the TLS survey and the photogrammetric bundle adjustment. 3D coordinates from each of these total station surveys were then derived.

These two sets of coordinates were then mapped onto each other using a least squares solution in CloudCompare, which minimized the Euclidean distance between both sets. The RMS residual error was noted in each case (Table 3.6). This could be reduced by removing points, though a decision was made to utilise all of the GCPs within the bundle adjustment, owing to the increase of redundancy from using 8 points. These final point coordinates were used to transform the TLS survey onto the local coordinate grid, as well as used in the bundle adjustment stage within the photogrammetric software.

Table 3.6. Residual error from Total station registration

Site	TS RMS residual error
Hunstanton	5.36 mm
Overstrand	19.46 mm

3.4.2 Terrestrial laser scan survey

The terrestrial laser scan survey, using a Leica P40 laser scanner (Leica, 2017a) paralleled that of the total station. Two scans were performed at each location from either end of the transect. Sampling rates were estimated to be near the rate of at least the GSD of images within the photogrammetric survey.

Table 3.7. Summary of settings for TLS capture. Estimated error is from the manufacturer documentation (Leica, 2017a)

Site	Sampling distance	2D Sampling rate (XY)	Estimated 1-sigma error (XYZ point)
Hunstanton	50 m	0.006 m	0.003 m
Overstrand	75 m	0.006 m	0.003 m

Once the scans from each position at each site were coregistered using Leica's Cyclone software package, they were exported as E57 point cloud files, which were subsequently converted to .LAS files, an open source file format. The unified clouds were then mapped onto the same coordinate system as the GCPs by performing a similarity transformation between the targets identified

within the TLS cloud and the coordinates derived from the Total Station survey. Thus, when it came to measure the cloud-to-cloud distance between the photogrammetrically derived point clouds and the TLS, no further data manipulation was required.

The multiple scans generated some redundancy in the TLS surveys, and act as a higher quality reference surface to compare with the produced photogrammetric cloud. High quality scans of the TLS registration targets were also acquired to ensure accurate co-registration of each of them to one another. The residual error is listed within the table below, with both sites showing a high quality registration.

Table 3.8. Summary of rigid body transformation error for each tilt and turn target for each site

Site	Target number	3D XYZ residual error
Hunstanton	1	0.001 m
Hunstanton	2	0.001 m
Hunstanton	3	0.001 m
Hunstanton	4	0.001 m
Overstrand	1	0.003 m
Overstrand	2	0.002 m
Overstrand	3	0.003 m

For the Overstrand site, target number 4 was unable to be scanned at high resolution from one end of the transect, presumably because it was too far away to be accurately captured. Thus, only three of the targets were used for the similarity transformation, which may contribute to a systematic error seen within the Overstrand results. In general, the expected error for both sites was low. The slight increase in TLS registration error at Overstrand was likely a result of sub-optimal environmental conditions and targets being farther from the scanner.

3.4.3 Photogrammetric survey

A transect was chosen in order to emulate a ‘weak’ viewing geometry, where the principal axes of the cameras do not overlap, and where systematic error is generally noted (James and Robson, 2014). Camera stations were marked at eight approximately equidistant locations along a transect orthogonal to the cliff faces, to provide an 80% overlap between images captured at adjacent stations. The 35 mm lens on the full frame Nikon D700 camera led to GSDs of 7 mm at Hunstanton and 12 mm at Overstrand, as noted previously.

The camera was set to capture images with ISO set to the lowest level available on the camera to keep noise to a minimum (200), using shutter speed priority so that the shutter speed could change based on the cameras automatic white balance.

The Nikon D700 was set to an aperture of f/8 at both sites, to ensure sufficient depth of field to capture everything in the scenes, including the landslip at the front of the Hunstanton site and the full depth of the landslide at the Overstrand site. It was noted that slight diffraction effects would be expected at this aperture, with the circle of confusion having a diameter 1.15 times the width of the pixel size, but it was deemed an acceptable level to ensure the scenes were perceptibly in focus (O'Connor et al., 2017). At f/8, with the focus set to 30 m (for Hunstanton), the depth of field ranges from 4.3 m in front of the camera lens to infinity, with a hyperfocal distance (The closest distance at which the lens can be focused while keeping objects to infinity perceptibly in focus) of 5.14 m.

Using the Nikon D700, sets of images were captured from each camera station (Table 3.9)

Table 3.9. Summary of acquired image sets

Image set	Description	Number of images
Orthogonal	'Weak' geometry set, equivalent to UAV nadir	1 / station
Multi-angular (every 15°)	'Strong' geometry set, convergent imaging which should reduce systematic error within the bundle adjustment (James and Robson, 2014)	5 / station
Repeat images	The same images taken at each angle from each station, for use in potential denoising algorithms (Hytti, 2006)	5 / station
Exposure differences	Images taken at each angle from each station, with shutter speed changing to under- and over-expose images from between -2EV and +2EV in 1EV increments (Debevec and Malik, 1997)	5 / station

These led to 1,000 images being acquired from the Nikon D700 at each site, 125 images from each station. The shutter speeds varied from between 1/1600 s to 1/100 s between the sites, across all exposures acquired.

3.4.4 Image processing

Before input into photogrammetric workflows, each of the RAW images required development to get them to the point of testing (Table 3.10).

Table 3.10 List of test sets generated for experiments

Test set	Algorithm
Compression level	A standard RAW-JPEG conversion for every image from the middle exposure for each camera station was undertaken using ImageMagick. The 'quality' keyword was varied in the command line application, with nominal values of 10,25,50,75 and 92 (default) chosen to sample the compression space, representing compression ratios (filesize vs. that of the RAW file) of approximately 75:1, 35:1, 20:1, 12:1 and 4:1 respectively compared with the original images.
Greyscale	Raster calculator operations were performed across a search space inclusive of each of the three colour bands. This used the 'fx' keyword in ImageMagick, which allows numeric operations to be performed without low level processing required.
Geometry	Sets of images were developed for each of the imaging configurations presented.
Downsampling	Sets of images were downsampled in order to simulate data acquisition using a smaller sensor of 5/6ths, 2/3rds and 1/2 the size of a full frame sensor by area.
Noise	Sets of TIFs were generated and noise added to each with a specified SNR, assuming the original images were noise-free. This was performed using a custom Python script, with variance used to initialize the SNR calculation using the equations described in the previous chapter. SNRs of 18,21,24,27 and 30 were chosen, simulating a doubling of noise with each image set, starting from a near 'excellent image quality' (32 dB, ISO 2006) to just below an 'acceptable image quality' (20 dB, ISO 2006)

3.4.4.1 Additional greyscaling algorithms

As discussed within the literature review, a range of non-linear greyscaling algorithms were considered for inclusion within this contribution. Owing to previous research (O'Connor et al., 2016), two were shortlisted for inclusion in this work. Many of the algorithms from Table 2.1 are already included within the linear-combination analysis (Table 3.10). Other greyscaling algorithms (such as Grundland et al., 2007) produced highly variable results for single images, which reduced the quality of the registration in pilot studies.

3.4.4.1.1 PCA

Principal components analysis (PCA) is a well-documented technique for dimensionality reduction, which involves rotating a dataset to maximize variance along a given axis. For this technique, pixels are treated independently of one another, so any spatial trends are not preserved. While some have recommended against the use of PCA owing to its lack of ability to identify obvious changes in colour contrast in certain contexts, such as for print media (Rasche et al., 2005). Notwithstanding, PCA is a compelling technique as the first principal component (PC) does maximizes the statistical variation between each band when generating a single channel image using a linear combination of these channels.

It should be noted that PCA is essentially a linear combination of each of the red, green and blue channels, which is similar to other greyscaling experiments performed in the chapter. The reason PCA is different, is that these combinations are allowed to vary between images, so the final set is a combination of different weightings of red, green and blue channels depending on the statistical properties of each image.

3.4.4.1.2 PCALab

This technique was adapted from Verhoeven et al. (2015) and featured in O'Connor et al. (2016). It involves a colourspace transformation from RGB to Lab, followed by a PCA. This is known as image 1. Three further weighted RGB – Lab transformations are undertaken with weights of 0.25, 0.45 and 0.65 assigned to the Luminosity channel in order to maximize any residual colour information not captured from the first technique. A three-band image is subsequently generated from this, and the first PC of this extracted to give image 2. Image 1 and 2 are subsequently averaged to give the result (Appendix 3). It gave promising results in both previous research contributions.

3.4.4.2 Compression

The JPEG image format involves compressing 8x8 windows of pixels systematically using a discrete cosine transformation. Depending on a variable, known as ‘quality’, the user can control how aggressively images are compressed. This can be thought of as how many degrees of freedom the fitted function must model the 8x8 pixel input – with higher levels of compression the order of the cosine transformation is lower.

Posterization is an artefact associated with JPEG compression resulting from generalization of these pixel blocks containing similar tones. This creates ‘blocky’ transitions between windows with similar tonal components. Posterization is visible within images at higher levels of compression (Figure 3.6). The effect of compression is more acute in pixel blocks which have

internally high contrast, such as where there are large depth discontinuities, or where texture changes, and so localized differences would be expected within these areas when comparing point clouds.

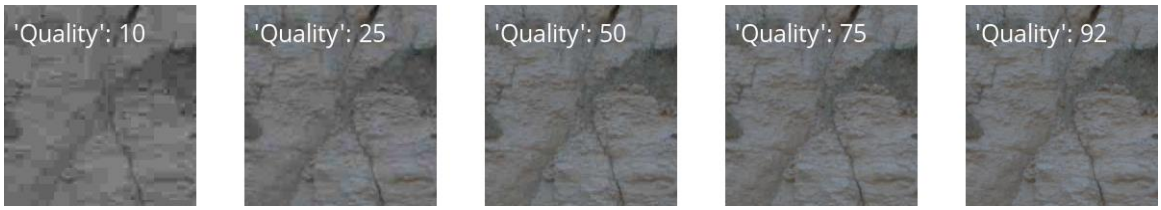

Figure 3.6. 200 x 200 pixel crop sections from JPEG images from the Hunstanton site of varying 'quality'. Decreasing 'quality' will apply more aggressive compression.

3.4.4.3 Noise

Noise levels in images from COTS cameras captured with the same settings can vary. Noise was added incrementally using the noise definition from DxOMark, (2018) in order to standardise reported additive noise with their reported metrics. For this experiment, the input images were assumed to be noise free. Whilst this assumption is incorrect, for the purposes of the experiments allows investigation of relative levels of degradation.

SNRs at levels of 18, 21, 24, 27 and 30 dB were added to each image within the original set to form 'noisy' image sets, with lower SNRs indicating more noise present. 200 x 200 pixel crop sections of these are presented in Figure 3.7, showing an increase in noise and its effects on how the image contents are perceived. At higher noise levels, such as 18 dB, which ISO classifies as below 'acceptable quality', fine details, such as the edges of the grass, are harder to perceive.

Figure 3.7. 200 x 200 pixel crop sections from JPEG images from the Hunstanton site of varying noise level within the images

3.4.4.4 Downsampling

Resizing, or downsampling, is a technique which has been used previously in order to reduce the computational demand of photogrammetric sets, though was associated with earlier studies and is now uncommon (Westoby et al., 2012). Regardless, no literature exists detailing the effects that

these downsampling operations have on photogrammetric outputs. Images were resampled using the corresponding TIF image after the downsampling algorithm had been applied, thus any effects caused by the algorithms acting in conjunction with one another would also be included within the result set. The 12 MP full resolution images were resampled to 91 %, 81.5 % and 70.5 %, to represent images containing approximately 10, 8 and 6 MPs respectively. Visualizing these results in a small 200 x 200 pixel crop Section of the cliff (Figure 3.8) shows the effects of the downsampling process, whereby sharp edges are generalized and fine detail lost.

Considering GSD is a commonly used predictor for survey accuracy (Nocerino et al., 2014), the intention was to measure the accuracy of the resampled sets to establish whether the magnitude of the reduction in accuracy was to do with the size of the pixel GSD, or the contents of the image, or both.

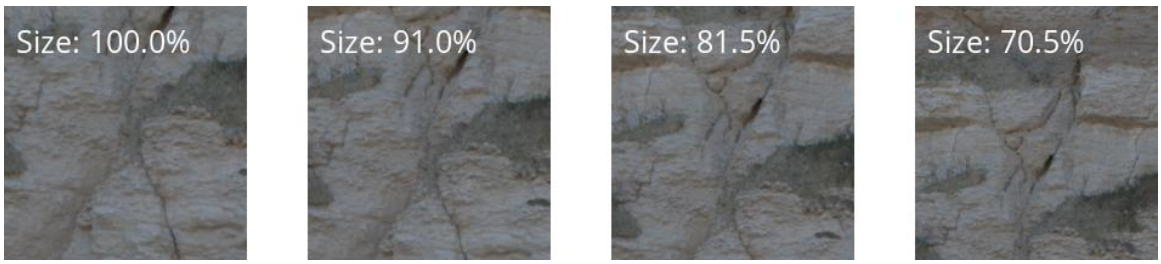

Figure 3.8. 200 x 200 pixel crop sections from JPEG images from the Hunstanton site of varying size of the image.

3.4.5 Photogrammetric processing

A workflow was developed to process each image set being tested. This involved standardising settings and processing across the entirety of the software packages utilised to perform the photogrammetric and point cloud post-processing.

For photogrammetric processing, Agisoft PhotoScan version 1.3.1.4030 (Agisoft, 2018) was used for both sparse cloud generation (ie. bundle adjustment) and dense cloud generation. A limit on the number of interest points which can be detected within an image, as well as the total number of possible points used to generate a matched image set can be controlled in the software package. For these experiments these limits were unconstrained. Images were not resized and/or compressed prior to use within the software package, unless part of an experiment. A Python script was developed to standardise this process (Appendix 1).

For the lens model fitted during the optimisation of cameras after the initial bundle adjustment, only one radial distortion coefficients (k1), principal point (cx, cy) and focal length parameters (fx, fy) were fitted after investigating correlations within the camera self-calibration (James et al.,

2017a). The second and third radial distortion parameters (k2, k3) were omitted from the bundle adjustment as they were highly correlated with the first radial distortion coefficient, the majority (k2 and k3 have correlation coefficients of -0.96 and -0.98 when compared with k1) of the corrections for radial distortion are achieved by using this single coefficient and removing the other two parameters reduced the risk of overfitting (Table 3.11). Tangential distortion coefficients (p1, p2) were also omitted as they are strongly correlated with the corresponding principal point parameters for the same reasons (cx and p1 have a correlation coefficient of 0.94, cy and p2 have a correlation coefficient of 0.9). This led to just four parameters being fitted for the Hunstanton survey (Table 3.12). For the Overstrand site, these same four parameters were used in bundle adjustment to initialize the cameras in the first instance.

Table 3.11. Initial correlation coefficients for camera calibration for the Hunstanton survey

	Value	**f**	**cx**	**cy**	**k1**	**k2**	**k3**	**p1**	**p2**
f	4255.41	1	0.03	-0.3	0.08	0.07	-0.05	0.02	-0.19
cx	-5.92213		1	0.02	-0.01	0.01	-0.02	0.94	0.02
cy	4.07361			1	-0.08	0.05	-0.07	0.02	0.9
k1	-0.087				1	-0.96	0.91	0.02	-0.01
k2	0.097					1	-0.98	0	-0.01
k3	0.008						1	-0.01	0
p1	0.00013							1	0.03
p2	-0.00037								1

Table 3.12 Correlation coefficients after removal of highly correlated values for the Hunstanton survey

	Value	**f**	**cx**	**cy**	**k1**
f	4255.41	1	0.05	-0.3	0.85
cx	-5.92213		1	0.02	0.03
cy	4.07361			1	-0.28
k1	-0.087				1

These parameters are detailed in full within the original paper describing Brown's camera model (Brown, 1966). The tie point and marker accuracies were set to values initialized after the processing of a high quality TIF set iteratively until the values converged (James et al., 2017a).

These were included with no limits set on the number of tie points the images could initially match on (Table 3.13, 3.14).

Table 3.13. Hunstanton photogrammetric parameters

Parameter	Value
Tie point limit	No limit
Sparse cloud quality	Highest
Marker accuracy	0.12 pixels
Tie point accuracy	0.27 pixels
Dense cloud quality	Ultra High
Dense cloud limit	No limit

Table 3.14. Overstrand photogrammetric parameters

Parameter	Value
Tie point limit	No limit
Sparse cloud quality	Highest
Marker accuracy	0.53 pixels
Tie point accuracy	0.43 pixels
Dense cloud quality	Ultra High
Dense cloud limit	No limit

For the blocks which were processed as part of the greyscaling experiments, the lens model could vary freely. This was to allow the photogrammetric processing to account for minor changes in focal length associated with each band in an RGB image. The lens models which were fitted for each set of greyscale experiments were subsequently exported for further analysis. Specifically, focal lengths in both the X and Y directions, which are fitted during photogrammetric processing, were plotted in XY space to establish the degree to which the software was compensating for changes in focal length due to the input channel.

Whilst no difficulties were encountered during the processing of the Hunstanton blocks, it proved difficult to produce consistent results from the Overstrand data. Initially, no camera positions were used to initialize the Overstrand site, but image matching and bundle adjustment were returning a range of gross errors during image matching, with very few frames registered. This was the case with self-calibration (no camera model used) also. Intervention was therefore

required to ensure image matching was stable for the Overstrand site. This involved performing image matching and bundle adjustment on the highest quality uncompressed TIF images 50 times. The block with the lowest pixel matching error then had its camera positions saved to an XML file. These were then used to initialize the positions of the cameras in the scene, ensuring a stable solution was returned.

Dense clouds were then systematically produced at the highest quality the software would allow. Once generated, the dense clouds from each block were exported as LAS files on a local grid for downstream processing. Due to the use of the GCPs within the bundle adjustment, the exported point clouds were in the same coordinate system as both the GCPs and registered TLS scans, allowing error metrics to be generated comparing all three.

Lastly, all clouds were clipped to the region of interest (the region of the scene covered by the TLS at the highest accuracy) to ensure results generated were only for areas covered by both the TLS and photogrammetric surveys. Inclusion of points outside of this area would bias results, as clouds with high quality images could potentially generate a significant proportion of points far from the TLS data, which would reduce their accuracy in the final experiments. This workflow is summarised in Figure 3.9, which is utilised throughout this work.

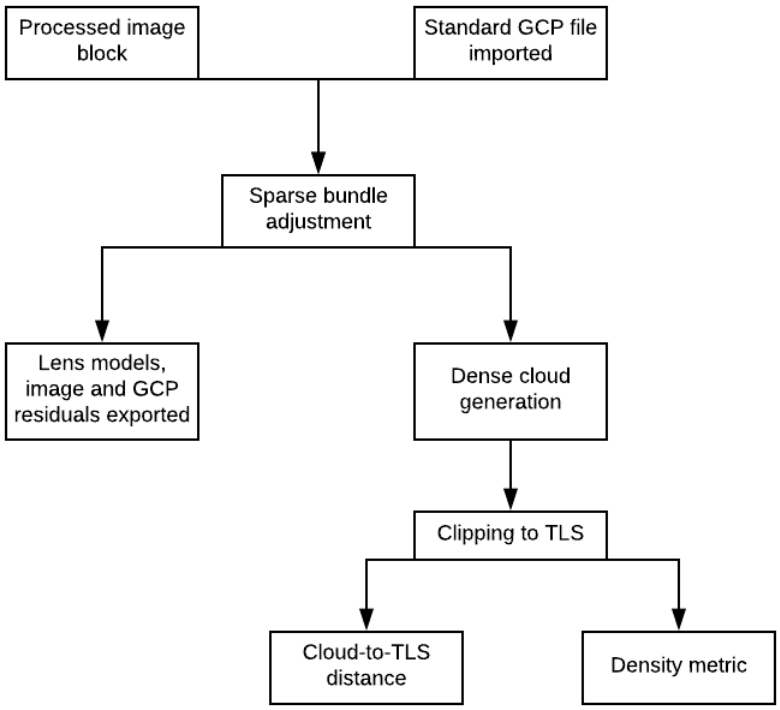

Figure 3.9. Standard workflow for photogrammetric tests.

3.4.6 Error metrics

3.4.6.1 Median error

The principal error metric used to derive quality estimates for each cloud produced through the photogrammetric software was the cloud-to-cloud distance (also known as the Chamfer matching distance (Barrow et al., 1977)), which takes each point within a cloud and finds the point with the smallest Euclidean distance within the reference cloud. This was calculated using the CloudCompare software package (version 2.8.beta). This is a commonly used metric within this kind of study as is shown by its frequency of use in the literature (Gómez-Gutiérrez et al., 2014; Micheletti et al., 2015). A maximum distance of 0.2 m was set between clouds in Chamfer matching, as outliers present in the photogrammetrically derived clouds slowed down the cloud-cloud matching. Points farther than 0.2 m away were thus clipped to this distance, with everything at or above this value being marked as a gross outlier.

As well as the threshold placed on the Chamfer matching, a further maximum distance between clouds of 0.03 m for the Hunstanton site and 0.2 m for the Overstrand site were set in order to ensure noisy points didn't overwhelm the signal within the clouds which had been generated. This was particularly important for the results from the Overstrand site, where some gaps appeared in the TLS point cloud in areas where vegetation was present. Where these areas were recovered or

partially recovered within the photogrammetric point clouds, these points would largely be marked as outliers as they are relatively far from the TLS cloud. In addition, points around the edges of these areas would be close enough to the TLS cloud to not be marked as outliers but would reduce the error inadvertently as the related reference point didn't exist. These last outliers were difficult to effectively deal with during cloud comparisons and are to some degree still present within the results. The vegetation was likely a large contributing factor to the accuracy being much lower at Overstrand. Whilst filtering the vegetation could be done to exclude these errors, it was not filtered in these experiments as it comprised most of the land cover at the site. The median error, rather than the mean, was also used in order not to bias outliers. After applying the Chamfer matching algorithm, the final point count, as well as error statistics associated with the clouds were noted before proceeding to processing the results.

3.4.6.2 Point count

The final number of points returned by the process for each cloud served as a second check on the quality of the product. While median error (ME) is currently the standard in the literature for comparison of results for a given survey, between clouds it serves to ensure that points aren't filtered out erroneously and only very high-quality points are kept. ME therefore needs to be presented in the context of the relative number of points a cloud produced to ensure this is not the case. When analysing results, clouds with more points but a higher ME are often including points which are less certain, but returning some value rather than a "no data" are generally desired behaviour.

3.4.6.3 Point density

Point cloud density is a metric useful for analysing how consistently points are spread across a cloud. High point count with relatively low density indicates that points are spread out evenly across a point cloud. Conversely, low point count with relatively high density would indicate points are mostly clustered in one part of the cloud, and there are likely gaps in coverage within the point cloud.

3.4.6.4 Pixel matching error (RMSE-P)

During the first stage of SfM processing (image matching) some error is present when images are mapped onto one another. This error represents the residual distance between feature points detected on each image when mapped onto one another. For these analyses, this was reduced to the root-mean-square error in units of pixels (RMSE-P) for points in the sparse point cloud, to allow for ease of comparison between the many results generated in these experiments. As with ME, it should be noted that lower error in this metric does not necessarily translate to better performance in the final point cloud product. It should be viewed in the context of both the ME

and point count – where both metrics are of high quality, and this error is also low, confidence is added to the quality of the product.

3.4.6.5 Difference of DSMs

Point clouds were sampled onto a raster grid at 0.02 m spacing to compare the final Digital Surface Models (DSMs). The raster grids average the values of all points projected onto the grid spacing for a given pixel on the grid, to create a more general spatial picture of where residual error lies. These grids were generated systematically using the CloudCompare software package, owing to its efficient memory use for these operations and the ability to standardise the process programmatically. These DSMs can then be directly compared visually, to analyse any spatial patterns which may be present between similar clouds, or DSMs subtracted from one another (difference of DSMs, such as in Westoby et al. 2012).

Chapter 4: Pilot studies and proofs of concept

In Chapter 3, two datasets are described which acted as pilot studies for the work done during and after the field work. These pilot studies sought to establish proofs of concept for two research exercises to be performed on the data from the field sites. The first is that degrading image quality will lead to a degradation in the quality of the SfM products, this was established by degrading the laboratory acquired dataset through increasingly higher levels of JPEG compression, processing it through the standard workflow (Figure 3.9) and comparing the results. Secondly, both the Westoby et al. (2012) dataset and the Hunstanton dataset were processed through the standard workflow, with image quality and SfM parameters kept constant, to measure the variability of the results of the standard workflow due to stochastic factors.

4.1 Results from varying image quality of lab data

The laboratory results were used to establish whether a difference between image sets using different compression levels could be ascertained. To do this, three subsets of images were generated with varying JPEG compression levels (25, 50 and 92) and compared with the reference dataset obtained from using the Faro Edge scanner (Figure 4.1). These compression levels were chosen, as they represent close to the worst case scenario, compression 25, which would only be used for images used for websites/applications. Compression level 50 represents the typical highest compression available off-camera (Nikon's 'basic' quality setting), and compression level 92 represents the typical default in software applications.

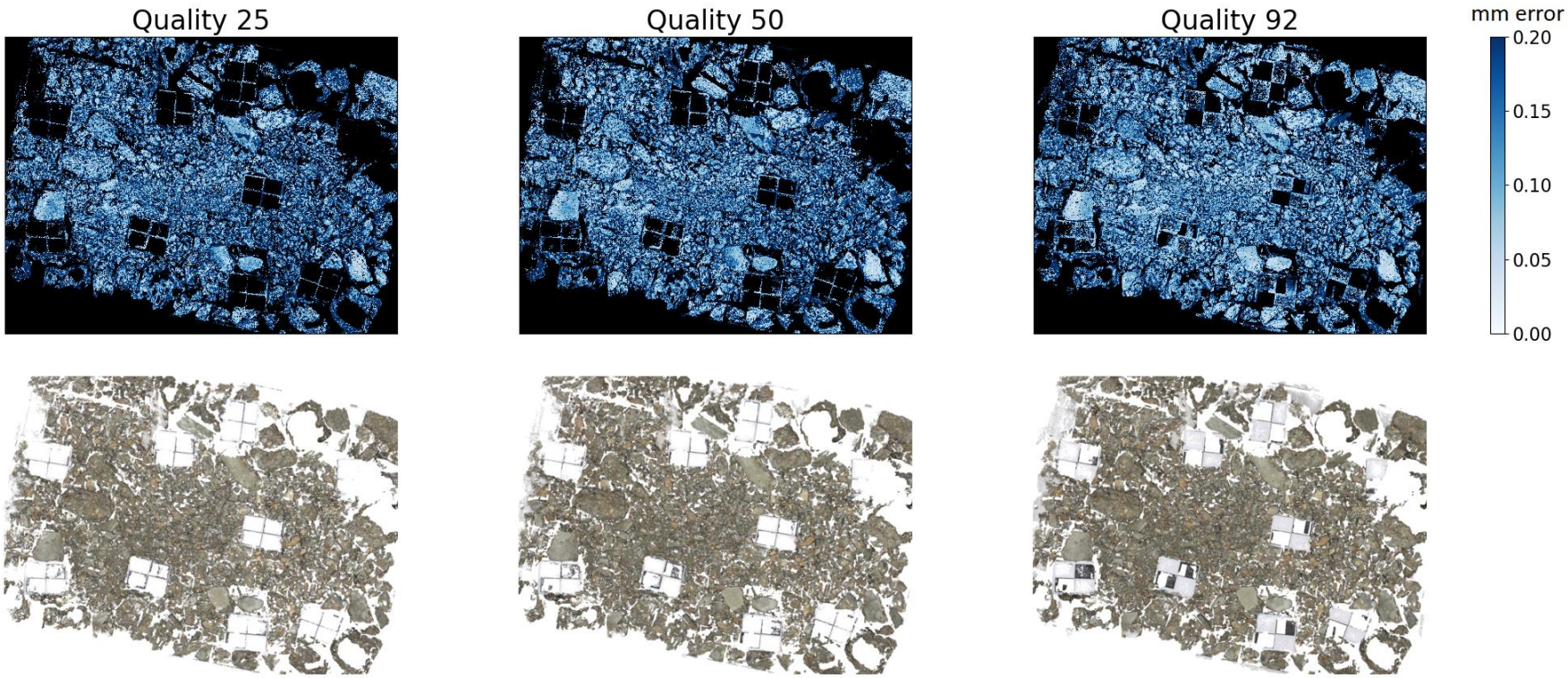

Figure 4.1 Cloud to cloud residual error of lab laser scan vs. compression level. RGB tifs are shown in the second row. Higher compression leads to higher error in the produced point cloud when compared with a high quality reference surface.

Visually the three clouds look quite similar, with notable differences being a lack of detail on the checkerboard targets on the two datasets of higher levels of compression when compared with those of lower compression. Table 4.1 shows a trend with degradation when the three image sets are compared, with the cloud-to-cloud median error increasing as the compression levels increase. Alongside this, a decrease in point count is seen with increasing compression. This is expected as the images become more compressed and data are lost.

Table 4.1 Results from laboratory compression experiment

Image set	Median error (mm)	Point count (millions)
'Quality' 25	.146	9.78
'Quality' 50	.140	10.77
'Quality' 92	.127	12.00

4.2 Results of stochastic effects on point cloud accuracy

As SfM is a stochastic process, this exercise sought to estimate the variation in point clouds which are produced, given the same input, through the standard workflow. The rationale for doing these experiments is to establish the degree the stochasticity effects the error metrics presented, to establish how consistent results are given the same input with the same parameters.

4.2.1 Constitution Hill

Constitution Hill data (Section 3.3.3) was used for initial investigations into variability due to stochastic effects. Significant errors (up to 32 cm median error) were seen within the results of this pilot study (Figure 4.2) due to extremely large gaps in the reference point cloud, as described within the paper (Westoby et al., 2012). However, the results on which the initial workflow was tested shows consistency when the same input is run within it for a given set of images. The objective was to establish how reproducible point clouds from any given dataset are for a given set of input conditions. For the five treatments done for this initial investigation, each set of images was processed through to dense cloud generation and comparison with the reference cloud, as they would be within the workflow used in the main body of this research.

For the JPEG RGB colour images sets, which were untreated before being used in the photogrammetric workflow, a range of 6 mm, which represents 1.9 % of the median value of 315 mm for the set of ten runs was obtained, with one run an outlier. For the other treatments, consistency is seen, though the green channel was noted to have a greater range than the other sets, representing 3.1 % of the median value (Figure 4.2).

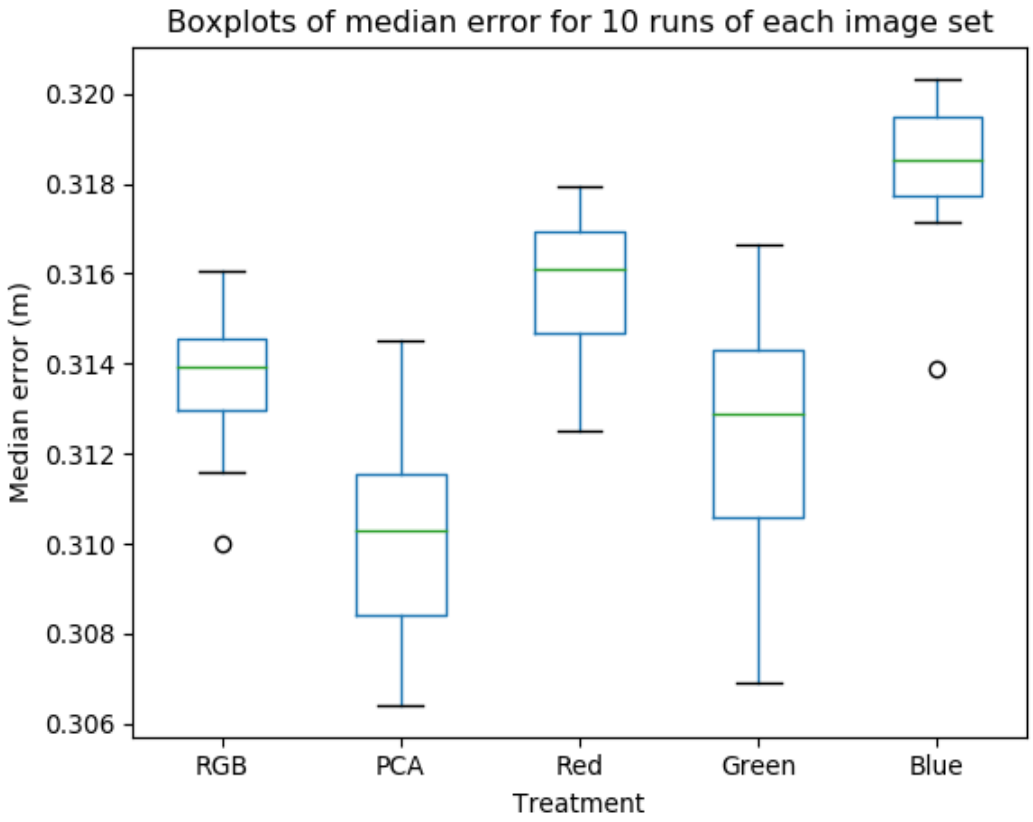

Figure 4.2. Variation seen within different treatments for desktop research using the Constitution hill dataset over multiple runs. Some consistency is seen within the results, with variation reaching at most 3.1 % of the median error. The green line represents the median of the sample, with boxplot whiskers representing90% of the full range.

4.2.2 Stochastic analysis: Hunstanton

A similar exercise was carried out for the Hunstanton dataset collected during the field work campaign. The full set of images was processed with the same settings through to the generation of errors statistics. Results show that for this higher quality dataset with large amounts of overlap, separate runs produced near identical results, showing a range of median error varying by less than 1/100 th of a millimetre. A minimum of 6.223 mm median error was obtained with a 6.232 mm maximum median error over 10 runs of the same dataset, indicating the results for the

Hunstanton site were extremely stable (Figure 4.3).

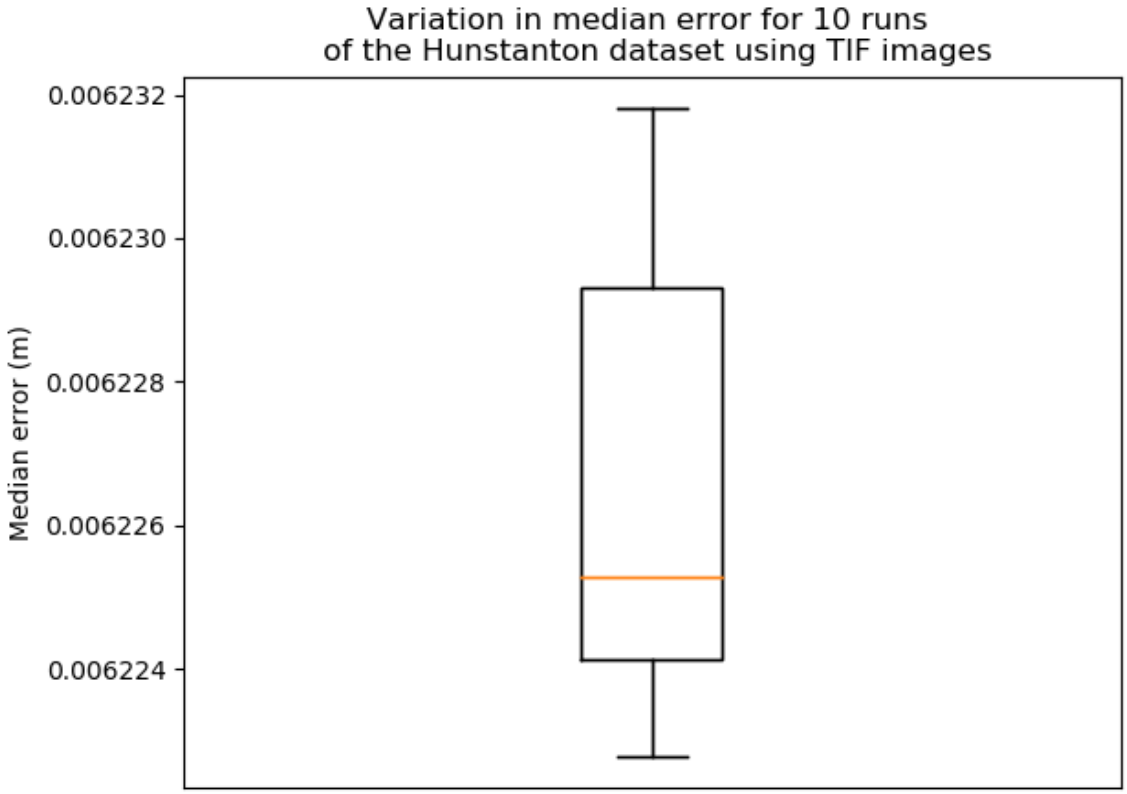

Figure 4.3 Results from processing the same RAW-derived TIF images acquired at Hunstanton through the full workflow show extremely similar results using the full set of images. The range represents 0.14 % of the mean error for the repeated runs.

4.3 Summary

The results of this chapter sought to act as proofs of concepts for the analysis to follow. Within this chapter, two main findings were presented, namely that image compression does influence the error statistics used to investigate point cloud quality, and that given a standard input, these results are reproducible. The latter finding is important for contextualizing the scale of difference between point clouds in the chapters to follow.

Chapter 5: Image quality

Within this chapter, results from the several different workflows for analysing image quality are presented. The experiments aim to establish whether there are interactions between the image quality and point cloud results. The standard workflow is utilised (Figure 3.9), along with the software for cloud-to-truth analysis. The JPEG compression level, size and amount of additive image noise of each image within each image set is varied to investigate how each impact the accuracy and density of photogrammetric point cloud products, against which the reference is compared. This is done by applying the various processing algorithms described in Section 3.3.4 to each image within the image set, and individually performing the photogrammetric processing on every set of images.

5.1 Photogrammetric results

5.1.1 Compression

5.1.1.1 Summary of statistics

For the Hunstanton survey, increased JPEG compression results in heavily degraded photogrammetric outputs (Table 5.1). The mean GCP error decreases with an increase in compression as the image matching and bundle adjustment give a higher weight to the GCPs within the processing steps. The mean GCP errors range from 3.6 mm for the most compressed image set, up to 5.41 mm for the quality 92 set. Conversely, an increase in RMSE-P is seen in the more heavily compressed sets of images. The median error for the dense clouds increases overall with an increase in this compression, ranging from a minimum of 6.22 mm in the uncompressed block to 9.44 mm the most compressed block. The point count is also seen to reduce dramatically in the image set with the highest compression, recovering less than half the points when compared with the uncompressed image set. A slight difference is seen in median error between the least compressed JPEG at quality 92 (6.49 mm) and the uncompressed image set (6.22 mm) representing an increase in median error of 4.3 %.

For the Overstrand survey a more ambiguous set of results was obtained than that from the Hunstanton site for compression results. The mean GCP residual does not appear to have a relationship with level of compression, though it is noted that the TIF set produced lower error than the default JPEG of ‘quality’ 92 (Table 5.1). While certain trends seemed to be similar, such as the RMSE-P increasing with increasing compression, and point count showing an increase with decreasing compression, the median cloud-to-cloud error did not seem to display the same trend as with the Hunstanton equivalent. The median error ranged from 17.7 mm to 19.6 mm, with the JPG image set of quality 92 showing the highest error of every block. The uncompressed images

returned the second highest error of all image sets, implying a worse performance in end to end products than that of the more compressed quality 50 image block.

Table 5.1. Summary of statistics from varying compression at the Hunstanton site. Quality is the level of JPEG compression applied, with a lower value implying increased JPEG compression.

Site	Quality	Mean GCP error (mm)	RMSE-P	Median error (mm)	Point count (millions)
Hunstanton	10	3.61	1.55	9.44	8.62
Hunstanton	25	4.86	0.71	7.73	16.33
Hunstanton	50	4.56	0.49	7.04	18.47
Hunstanton	75	4.52	0.39	6.7	19.4
Hunstanton	92	5.41	0.32	6.49	19.63
Hunstanton	TIF	5.36	0.28	6.22	19.84
Overstrand	10	7.86	1.98	19.40	9.82
Overstrand	25	10.33	1.17	18.30	18.83
Overstrand	50	8.37	0.82	17.70	20.52
Overstrand	75	11.52	0.59	18.50	21.38
Overstrand	92	10.00	0.76	19.60	22.02
Overstrand	TIF	8.39	0.39	19.40	22.20

5.1.1.2 Spatial variability

The DSMs produced from the point clouds show very slight variations between the image sets of compression 50, 92 and the uncompressed set at Hunstanton (Figure 5.1). Within the most compressed set in the comparison (‘quality’ 50) we can see a distinct increase in error in areas of fine detail. In terms of pixel occupancy, compressed results are complete with almost identical occupancy except for areas to the lower right of the cliff face. The area of the upper left of the cliff face is reconstructed with higher accuracy within the uncompressed image set. There is a distinct thinning out within this region in the highly compressed set, with pixels unoccupied or of densities less that of the uncompressed cloud image set. Within some areas to the left of the central landslip the ‘quality’ 92 image set produced areas of denser points when compared with the uncompressed image set.

Analysing the DSMs associated with the quality set to 50, 92 and uncompressed images from Overstrand reveals only very slight differences in occupancy of the DM cells and the spatial distribution of the error. Error tends to be concentrated in areas around the edges of patches of

grass, where the reconstruction contains an occlusion. Thus, the higher error is potentially linked to the ability of the uncompressed block to match more points within the dense algorithm which are in these areas, which could artificially inflate the error metric as more uncertain points are being included.

Almost no difference is seen in a textured area towards the back of the landslide between each set in terms of both median error (Figure 5.1 E) and point density (Figure 5.1 F). The only evidence of degradation is a very slight thinning is seen in the ‘quality’ 50 cloud in this area.

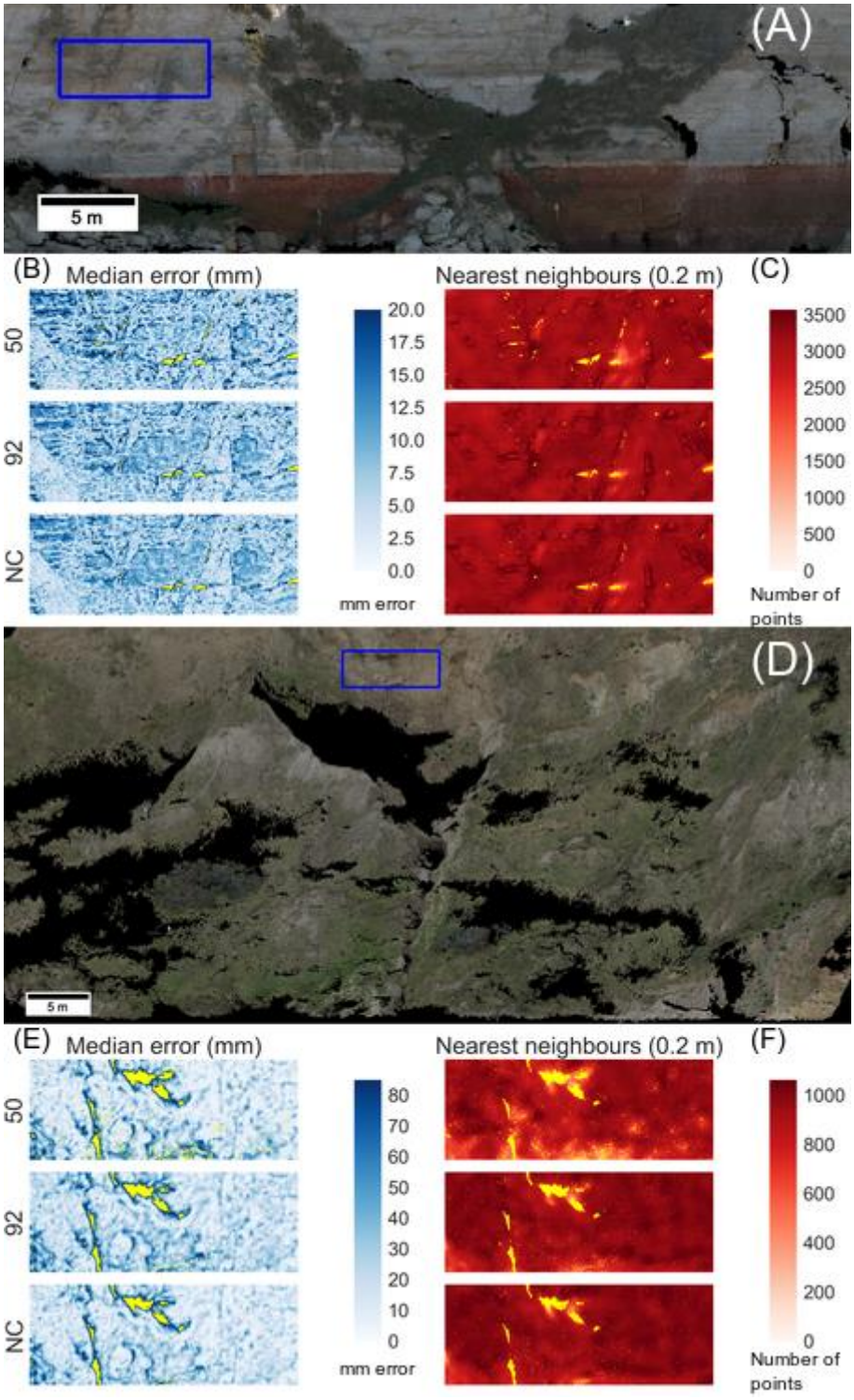

Figure 5.1. Hunstanton cliff shows degradation in cloud-to-truth error (B) and density (C) with varying jpeg compression, noted in the upper portion of the cliff face (shown in A). For the Overstrand results, very little difference in error is seen (E), even in a highly textured region located in the back portions of the landslide (shown in D). No data (missing) values are in yellow.

5.1.1.3 Distribution analysis

The median and quartile error increases according to how compressed the image block is for each of the quality settings for the Hunstanton results (Figure 5.2). For density, the inverse is seen to be true, with point count decreasing with increase in compression, as might be expected due to the reduction in data associated with these images.

For the Overstrand results, the violin plots reveal a very similar distribution when it comes to median error for each set of images, suggesting that the image quality is not the limiting factor when it comes to photogrammetric reconstruction. However, point density increases uniformly with decreasing compression, suggesting that the assertion that the points included around the edges of occluded areas could be contributing to the median error increasing.

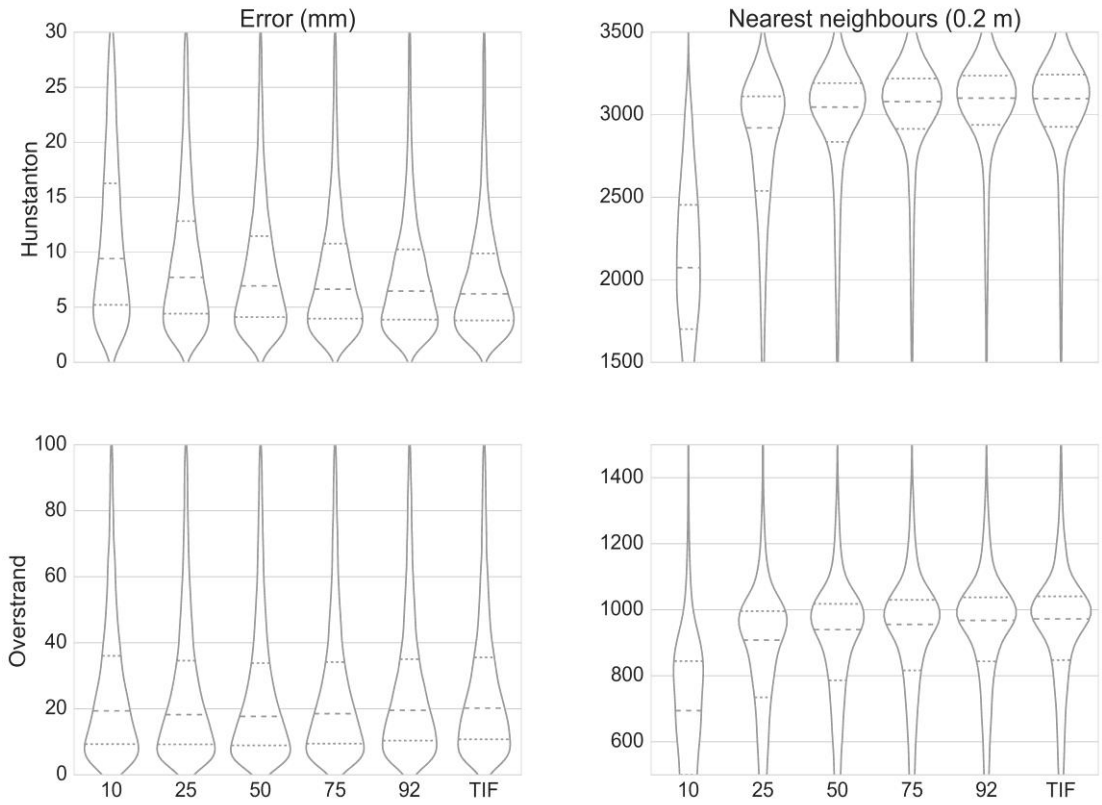

Figure 5.2. Median error increases with increasing compression for Hunstanton, while point density decreases. For the Overstrand results, point density decreases with increasing compression, though no trend with regards to error is seen. The dashed line indicates the median, with the dotted line representing the quartile ranges (25th and 75th percentile).

5.1.1.4 Difference of DSMs (DoD)

To ascertain whether any spatial trends exist within the cloud-to-cloud distances between two of the analysed clouds, difference of error DSMs were generated (as described in Section 3.4.6.5). These allow direct comparison of error DSMs on a pixel by pixel basis. Figure 5.3 shows the DoD between the error DSMs produced using images compressed using a 'quality' setting of 92, compared with those using a 'quality' setting of 50. Negative (blue pixel) values indicate where the higher quality images are performing better than the lower quality. This DoD demonstrates that the degradation to the photogrammetric product shows no strong spatial trends at Hunstanton, suggesting that the compression is in fact impacting the whole product, and not just one area.

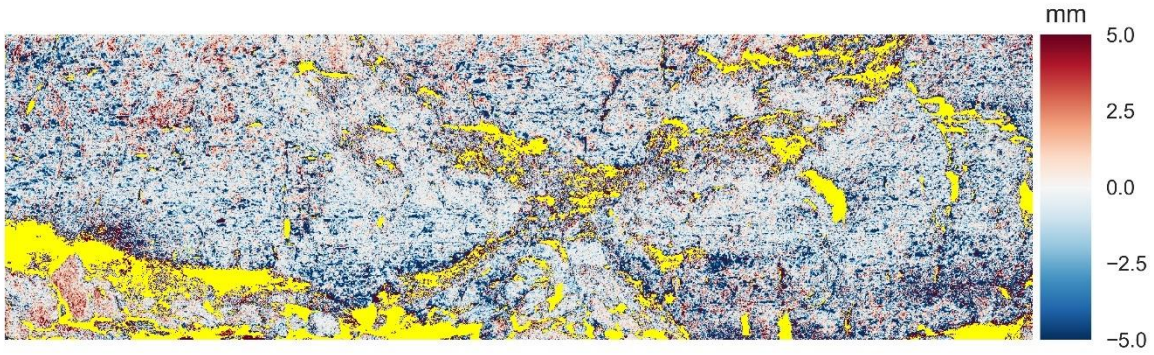

Figure 5.3. Difference of error DSM for the cloud produced from images compressed to 'quality' 92 compared with 'quality' 50 for the Hunstanton site. Blue areas are where the more heavily compressed block is underperforming relative to the less compressed block. No spatial trends are seen for degradation of accuracy due to compression. No data (missing) values are in yellow.

The Overstrand results don't show the same trend, with the higher quality images performing worse in terms of median error, as noted in Table 5.1. A vegetated region towards the back of the landslide was better reconstructed in the products from image sets with higher compression (Figure 5.4).

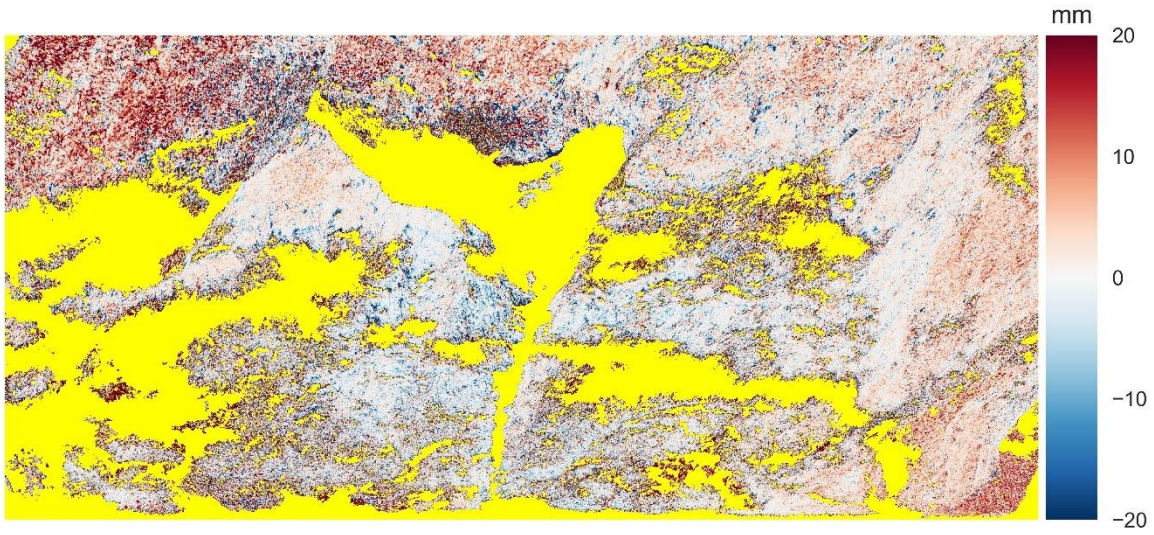

Figure 5.4. Difference of error DSM for the cloud produced from images compressed to 'quality' 92 compared with 'quality' 50 for the Overstrand site. Blue areas are where the more heavily compressed block is underperforming relative to the less compressed block. Higher levels of compression produced better results in terms of accuracy for occupied pixels. No data (missing) values are in yellow.

5.1.2 Noise

5.1.2.1 Summary of statistics

Adding noise to the image sets at Hunstanton shows a reduction in accuracy of the point clouds correlating to the amount of noise added to each image set (Table 5.2). GCPs were all closely related in terms of the reported error, though lower levels of error were associated with the point clouds with which more noise was associated, implying that the GCPs were more heavily relied on within the bundle adjustment. This trend was seen in every image block as noise was added, though only very slight changes were noted with an RMSE range of 5 – 5.3 mm. As more noise was added to images, the RMSE-P within image matching increased, with an RMSE-P ranging from 0.28 (uncompressed) to 0.66 (SNR 18 dB) suggesting a correlation between accuracy of the pixel matching and noise level. The level of noise also appeared to influence the number of points returned within each image set, with high noise levels returning less points. The range of median error was 6.22 mm (from the uncompressed set) to 7.37 mm (the SNR 18 dB set), representing a percentage difference of 18.5 %. The best performing block with regards to median error was the image block with no noise added. The cloud produced from the set with the highest level of additive noise (18 dB) had the highest median error and the largest RMSE-P, suggesting the degradation is reflected in the produced cloud.

Adding noise to the image sets at Overstrand had a similar effect to that of compression. In terms of the GCPs, increasing more additive noise led to an increase in the marker error during image matching, suggesting the software was relying more heavily on the GCPs to fit the model at higher levels of degradation (Table 5.2). Marker error decreased as noise increased in all but two cases (21 and 27 dB) where error was comparable to the next noisiest image block. As more noise was added to images, the RMSE-P within image matching increased, though this increase in matching accuracy was not reflected within the median error of the clouds produced. The level of noise influenced the point count, though not to the point where matching failed or holes would open up within the produced clouds. The range of median error was very narrow, between 20.2 and 20.59 mm. The best performing block in this regard was the uncompressed image block. The set with the least amount of additive noise performed the worst. The cloud produced from the set with the highest level of additive noise (18 dB) had the second highest error in terms of cloud-to-cloud, and the largest RMSE-P, suggesting the degradation is reflected in the cloud produced.

Table 5.2. Summary of statistics from varying noise at the Hunstanton site

Site	SNR (db)	Mean GCP error (mm)	RMSE-P	Median error (mm)	Point count (millions)
Hunstanton	18	5.01	0.66	7.37	16.4
Hunstanton	21	5.04	0.53	6.98	17.8
Hunstanton	24	5.11	0.45	6.73	18.9
Hunstanton	27	5.21	0.39	6.6	19.2
Hunstanton	30	5.24	0.35	6.45	19.5
Hunstanton	TIF	5.32	0.28	6.22	19.8
Overstrand	18	7.34	0.74	20.47	20.41
Overstrand	21	7.83	0.70	20.32	21.47
Overstrand	24	7.60	0.55	20.33	21.94
Overstrand	27	8.45	0.50	20.37	22.12
Overstrand	30	8.36	0.45	20.59	22.21
Overstrand	TIF	8.39	0.39	20.20	22.22

5.1.2.2 Spatial variability

At Hunstanton, minor spatial differences are seen with the error distribution for the DSMs produced. Pixel occupancy is very similar between each cloud, though evidence of some areas proving more difficult to reconstruct in the upper left portion of the cliff were seen (Figure 5.5 B).

Error is concentrated in areas which are typically more difficult to accurately reconstruct, such as those with large depth discontinuities and vegetated sections. Points retrieved within the landslip region in the centre of the DSMs show high error in all three examples. Within certain areas along the cliff face, however, the additive noise does appear to have degraded the quality of the product, such as the darker area to the lower right of the DSMs and areas where high frequency information is concentrated, such as areas of high contrast on the cliff face. This degradation is most clearly seen in terms of density, with the noisier image block showing reduced density in these portions of the cliff and a drop in pixel occupancy.

For the Overstrand results, error was concentrated in similar areas to Hunstanton, such as around the edges of where occlusions lie, and where points were retrieved in heavily vegetated or topographically complex areas, such as the middle of the gully. Density is highly correlated with distance from the imaging transect, with parts of the lower right of the landslide being more densely populated then others along the front. Densification is noted on the upper parts of the back of the landslide in both the uncompressed and SNR 30 clouds, with the SNR 18 cloud showing a thinning towards the upper end.

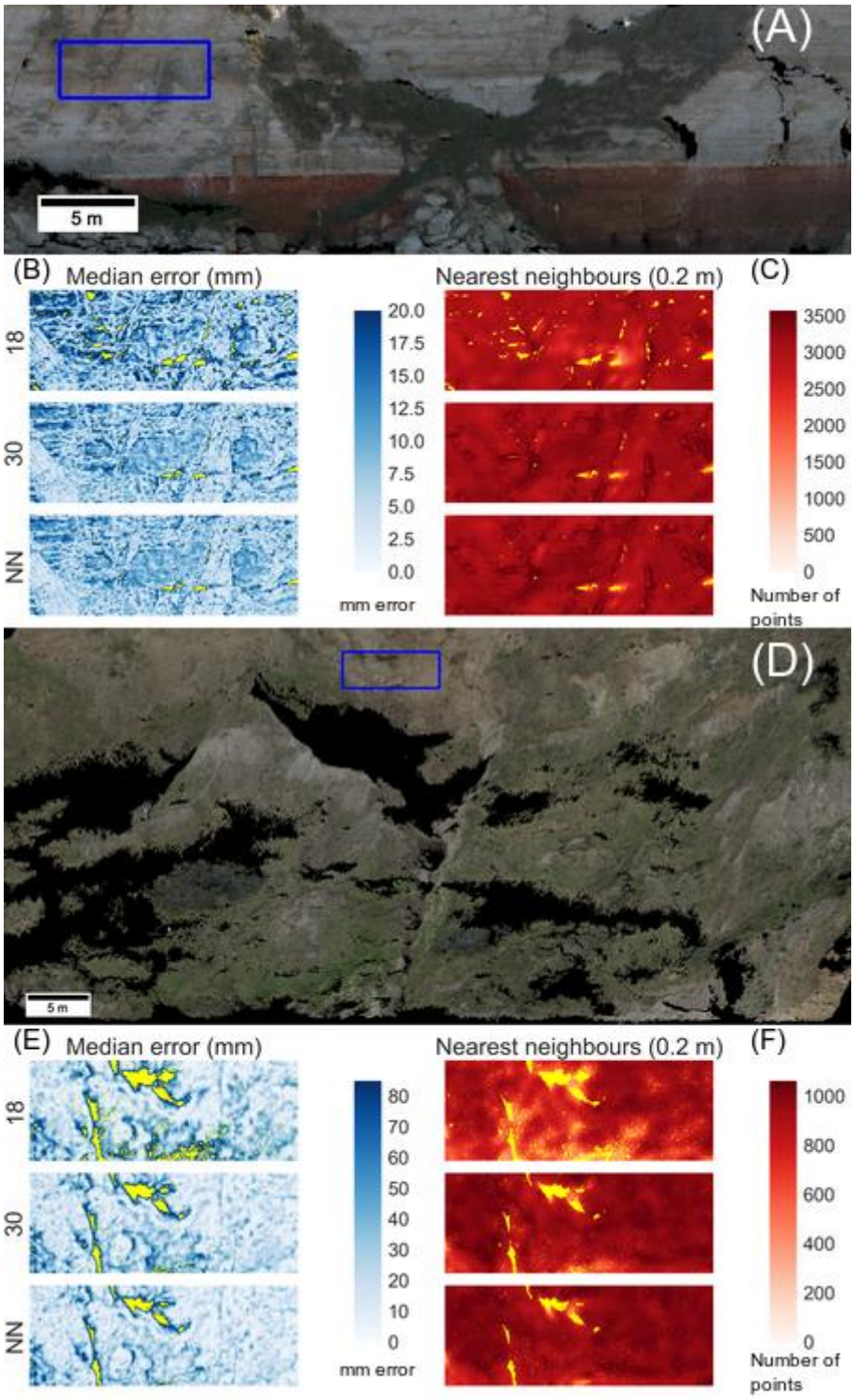

Figure 5.5. Effects of noise with introduction of no noise (NN) and additive noise of 30 dB and 18 dB. The Hunstanton cliff shows degradation in cloud-to-truth error (B) and density (C) with increased additive noise. The Overstrand landslide does not show these same trends, with a small Section of the highly textured back wall showing very little difference in error (D), though some thinning of the point cloud at high noise levels (F). No data (missing) values are in yellow.

5.1.2.3 Distribution analysis

In terms of error for the Hunstanton results, the quartile range and median error associated with each cloud increases with increasing noise, suggesting the image quality is the limiting factor with regards to error at this site (Figure 5.6). The same trend is seen within the density of each cloud, with median and both quartiles for every cloud decreasing with increasing noise, suggesting noise is impacting density in this case.

In terms of error at Overstrand, the quartile range and median error associated with each clouds remains more or less static, at a median of 20.3 mm for each image set. The additive noise does not show any signs of degrading these statistics, though at higher noise levels it is noted that point density does drop off, albeit at a low level.

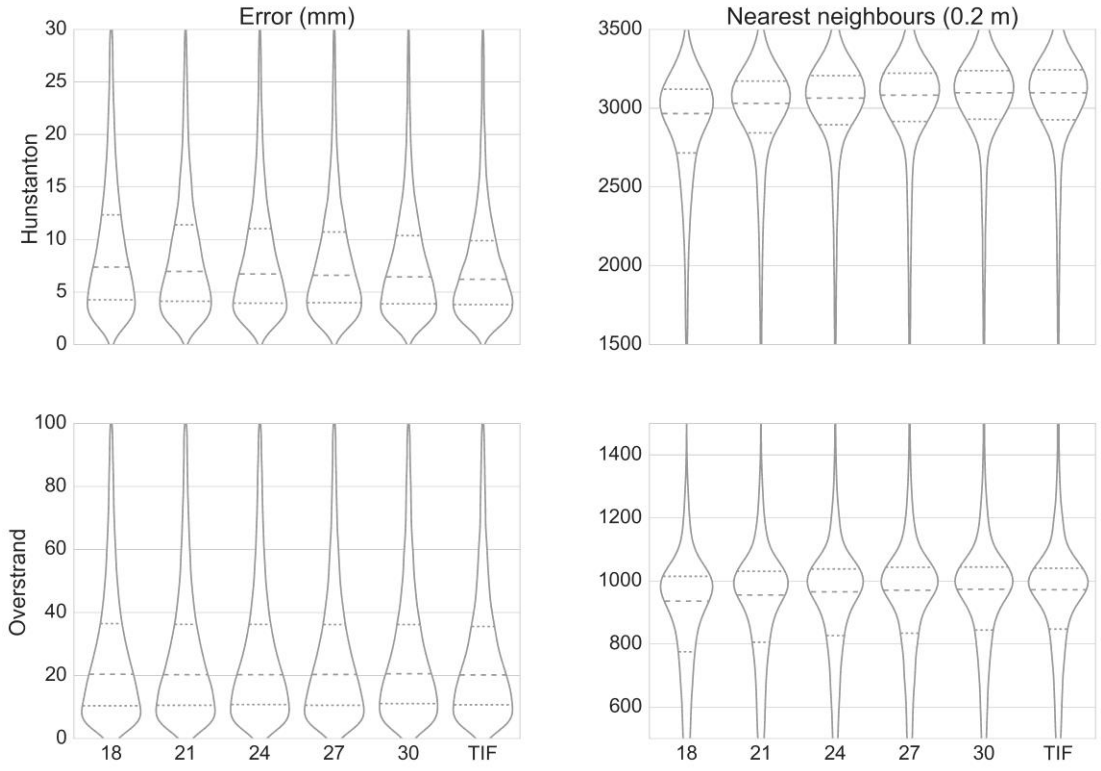

Figure 5.6. Median error increases with increasing noise, while point density decreases for the Hunstanton results. For the Overstrand results, no evidence of this trend is seen.

5.1.2.4 Difference of DSMs

The DoDs reveal that, like the DoDs from the compression sets, no apparent spatial trend exists between blocks containing various levels of noise at the Hunstanton site (Figure 5.7). The degradation (blue pixels) appears to be somewhat ubiquitous across the site, though is more pronounced towards the edges of the block (where less images generate the points) and where

the site is dark, where we would expect the noise to take up a bigger part of the signal (such as the bottom right region of Carstone).

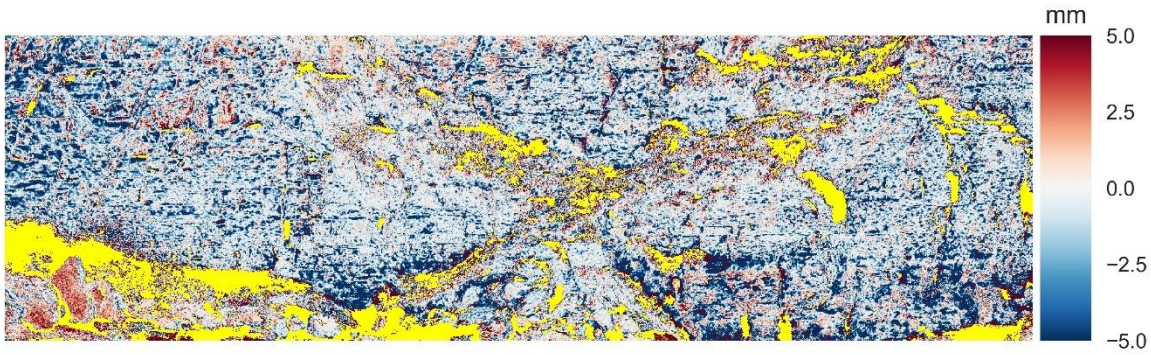

Figure 5.7. Difference of error DSM of the original image set compared with that with additive noise of 21 dB for the Hunstanton site. Blue regions are those where the original set performs better in terms of error. Increased noise degraded darker areas, such as the carstone to the bottom right, or shaded regions of the rock face to the upper left. No data (missing) values are in yellow.

At Overstrand, this is not the case, with the image noise appearing to show little to no spatial pattern (Figure 5.8). The trend of most of the difference being seen at the front right of the landslide and the back wall is seen across the result sets.

Figure 5.8. Difference of error DSM of the original image set compared with that with additive noise of 21 dB for the Overstrand site. Blue regions are those where the original set performs better in terms of error. No spatial trend was seen in terms of error with noise. No data (missing) values are in yellow.

5.1.3 Downsampling

5.1.3.1 Summary of statistics

GCPs were all closely related in terms of the residual error, with no apparent trend being shown in the range of image resizing presented (Table 5.3). The RMSE range for the GCPs for each cloud produced was 5.11 – 5.55 mm, though this could in part be explained by differences in marker localisation within the software, as for this stage markers needed to be located independently for each block with different image sizes. The RMSE-P per block tended to be lower in the smaller images, which suggests the downsampling is increasing the ability of the matcher to operate in terms of each pixel. In terms of point count, we see a very apparent trend of the number of points reducing with reduction in the size of the image, which is expected due to the lower number of samples the dense matcher can fit images to. The range of median error was 6.22 mm (from the uncompressed set) to 7.72 mm (the set resized to 70.5% of the original), representing a percentage difference of 24 %. A trend which showed a reduction in median error with increase of image size is seen within this experiment, suggesting a correlation between the size of the input image and the accuracy of the point cloud produced.

Resizing images at Overstrand showed no discernible trend with regards to GCP accuracy and median error between each of the point clouds. Within the RMSE-P residuals, a trend is seen whereby matching accuracy decreases with the smaller images as the pixel size gets bigger. This is not the case with the full sized TIF images, which have the smallest RMSE-P (0.39 pixels), which is almost half the RMSE-P of the full-size JPG image block (0.76 pixels). A similar trend is seen with the point count, with reductions as the size of each image is reduced, ranging from 10.61 70.5% image set) to 22.02 million points (full sized images). The minimum error seen within the clouds is within the clouds resized to 81.5 % of the original (19.10 mm).

Table 5.3. Summary of statistics from varying image size at the Hunstanton site. *indicates JPEG compressed image set

Site	Resize percent	Mean GCP error (mm)	RMSE-P	Median error (mm)	Point count (millions)
Hunstanton	70.5	5.50	0.30	7.70	9.50
Hunstanton	81.5	5.11	0.30	7.00	12.80
Hunstanton	91	5.55	0.31	6.80	16.10
Hunstanton	100*	5.41	0.32	6.49	19.30
Hunstanton	TIF	5.36	0.28	6.22	19.80
Overstrand	70.5	10.78	0.44	27.00	10.60
Overstrand	81.5	9.33	0.43	19.10	13.60
Overstrand	91	9.89	0.63	25.80	17.40
Overstrand	100*	10.00	0.76	19.60	22.00
Overstrand	TIF	8.40	0.39	20.20	22.20

5.1.3.2 Spatial variability

The DSMs produced show almost no difference in terms of pixel occupancy, except for some pixels surrounding the central landslip and the upper regions of the rock face (Figure 5.9 A-C). Error is again concentrated in similar areas, such as those with large depth discontinuities and vegetated sections. Density, however, is dramatically different, with median point density of the smallest images equating to approximately half that of the full sized uncompressed images. This is also approximately proportional to the reduction in the number of pixels per image, from 12 MPs (resulting in 19.8 million points) to 6 MPs (resulting in 10.6 million points). In terms of density of the image set resized from 12 MPs to 10 MPs (91 %), a drop in point count is seen from 19.8 million to 16.1 million points, which is again approximately equal to the reduction in pixel count (20 % less pixels) and the reduction in points vs the uncompressed set (23 % less points).

For the Overstrand results, the muddy cliff face shows larger error within the 91 % resized image block, which accounts for part of the large median error seen (Figure 5.9 D-F). The large median error associated with the 70.5 % image block seems to be associated with the patchiness and uncertainty of points concentrated in the lower left regions of the DSMs, where poor image overlap and dense vegetation were present, leading to very few points being produced in this area. Density DSMs demonstrate the trend seen in the table, with the clouds more dense towards

the front of the landslide, thinning out towards the back parts, with no discernible local variations between the resized image clouds.

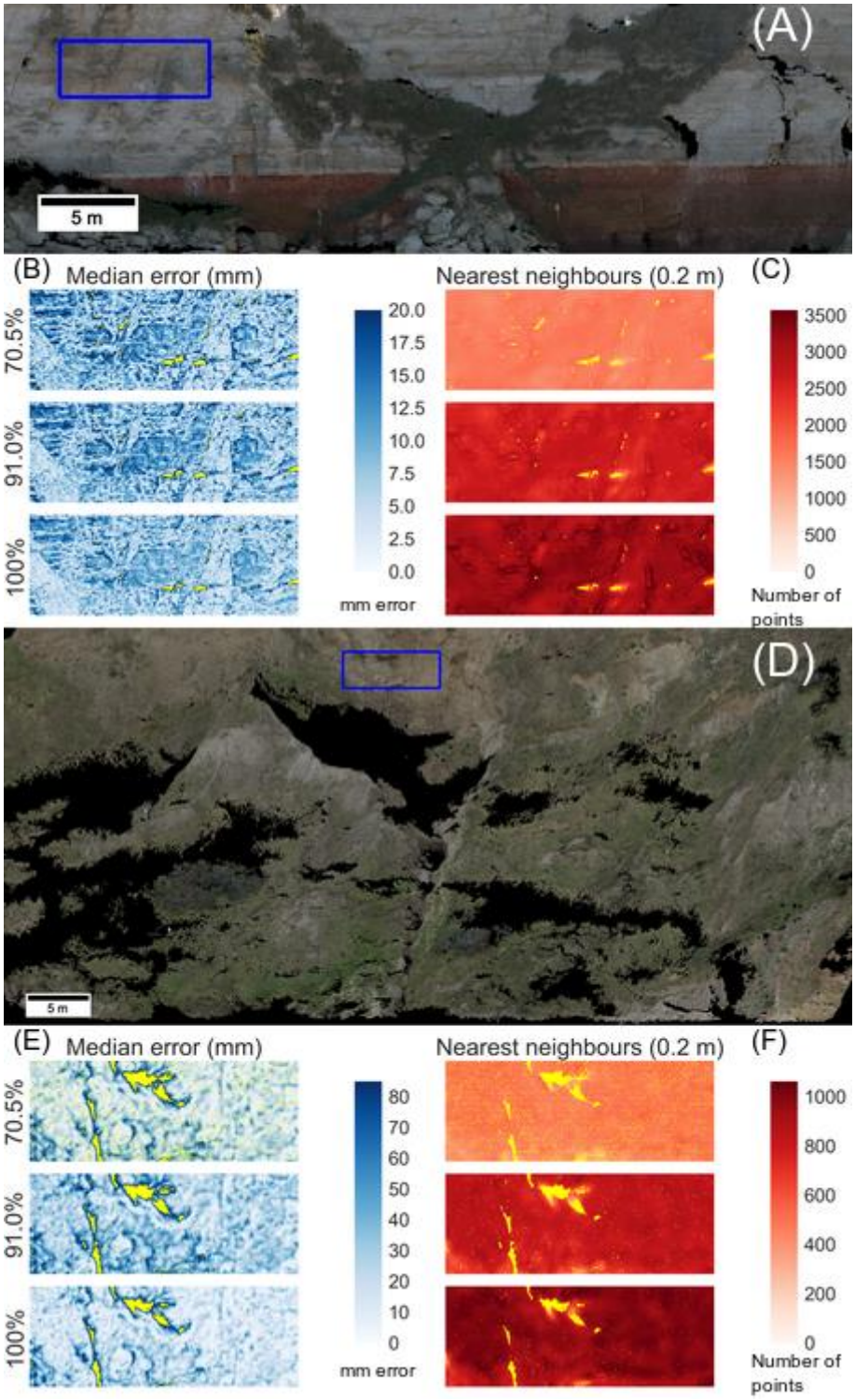

Figure 5.9. Hunstanton and Overstrand cliffs shows small degradations in cloud-to-truth error (show in and b and e) and large differences in density with varying image pixel count (c and f). No data (missing) values are in yellow.

5.1.3.3 Distribution analysis

For the Hunstanton results, the quartile range and median error associated with each cloud increases with decreasing resolution, with the distribution being consistent through the range of image downsampling (Figure 5.10). The density violin plot reveals the difference between resampled image sets, and the large contrast between downsampled point clouds.

In terms of error, the violin plots do not seem to produce any trend when it comes to the level the images were downsampled, with the minimum error coming from the image set resized to 81.5% of the originals. Density increases in an ordered manner as the images contain more pixels as with the Hunstanton site.

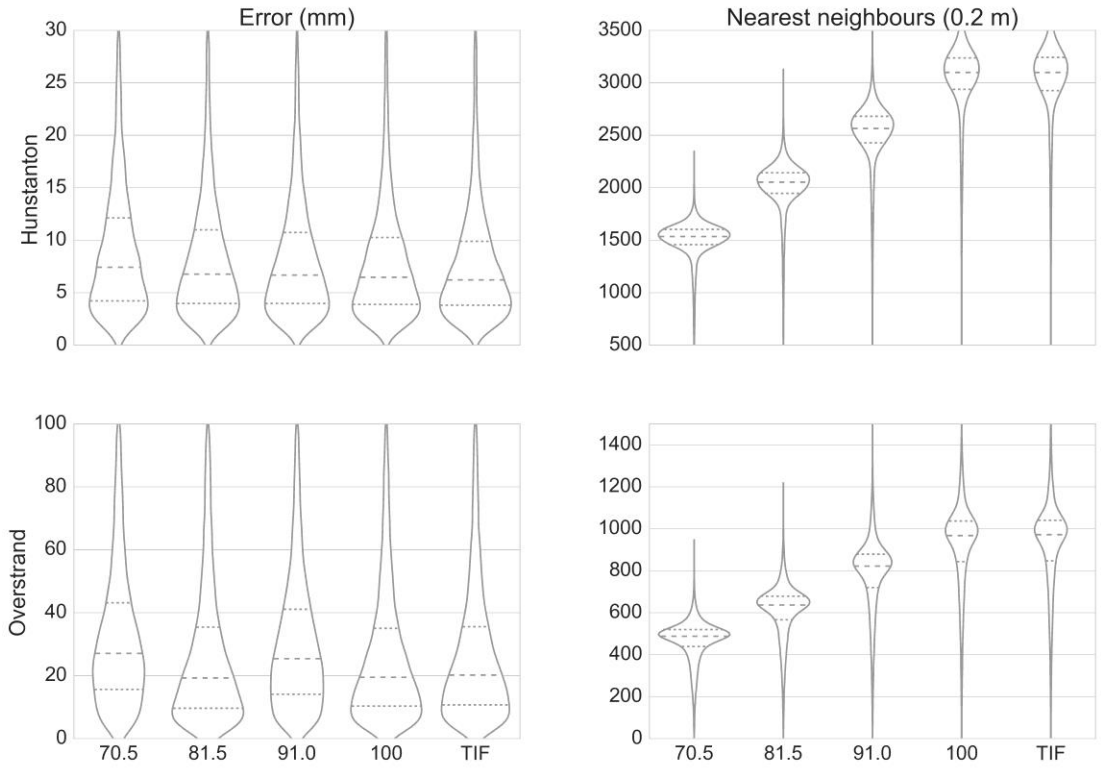

Figure 5.10. Median error increases slightly as images are downsampled, while point density decreases for both sites.

5.1.3.4 Difference of DSMs

As with both compressing and adding noise to the images, for the Hunstanton results the error appears to show no strong spatial trend (Figure 5.11), with error being lower in nearly every region of the DSM for the images with greater numbers of pixels. For the Overstrand results, the two sets seem to perform similarly (Figure 5.12), with some slight differences in the back Section of the landslide and the front wall towards the right of the figure.

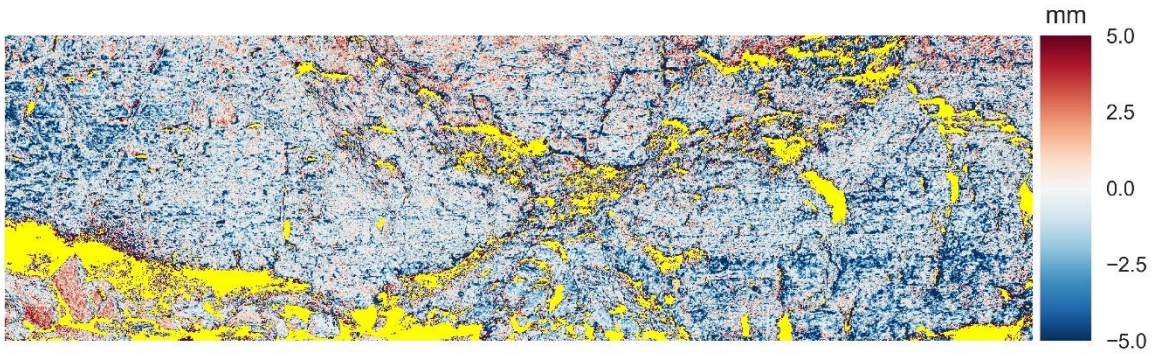

Figure 5.11. Difference of error DSM for the cloud produced from the original images compared with images downsampled to 70.5 % size for the Hunstanton site. Blue areas are where the downsampled block is underperforming relative to the original images. Downsampled images produced worse results in terms of accuracy at the Hunstanton site, with no obvious spatial trend observed. No data (missing) values are in yellow.

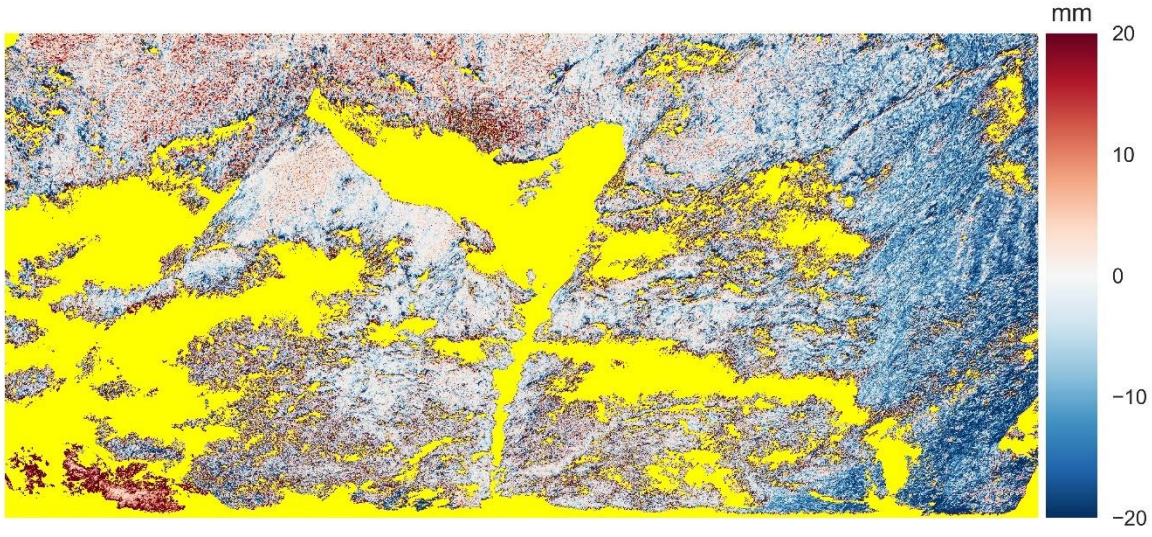

Figure 5.12. Difference of error DSM for the cloud produced from the original images compared with images downsampled to 70.5 % size for the Overstrand site. Blue areas are where the downsampled block is underperforming relative to the original images. The downsampled images

tended to return worse results in areas to the right of the landslide. No data (missing) values are in yellow.

5.2 Summary

5.2.1. Compression

The results at both sites suggest evidence of product degradation with increasing compression. Most notably, the trend of increasing compression decreasing point count when the settings used within the photogrammetric package were equal was seen. With regards to error, increasing compression led to decreased error at Hunstanton, though for the Overstrand results this trend was not seen. This suggests that the compression aided the image registration, and in turn the dense cloud generation. Even with the increase in pixel matching error between images of higher compression for the Overstrand results, this was not reflected within the final products. This is counter-intuitive, but in essence means the image processing is independent from the results, suggesting it is not the limiting factor at Overstrand.

5.2.2. Noise

Adding noise to images before processing had a slight negative effect with regards to the accuracy and density of the photogrammetric products for the Hunstanton survey. Worth noting is the scale of degradation with regards to the amount of noise – even with large amounts of noise added (18 dB), only a slight degradation was observed compared with the original image set for both sites. As with compression, increasing noise had more of influence over the quality of products at Hunstanton when compared with Overstrand.

5.2.3. Downsampling

Downsampling images had a dramatic effect on the point count of photogrammetric products at both sites. The degradation to the overall accuracy of the DSMs generated did not appear as acute, and while some degradation was seen for the Hunstanton results with increasingly downsampled images, no discernible pattern could be noted for the Overstrand results.

5.2.4 Insights and recommendations

The results from this chapter suggest that for geoscientific studies where lots of texture is present within the images and the scene is static (such as Hunstanton), maximizing image quality can led to higher quality results. For sites with complex geometry with contents that can move due to environmental conditions (such as the vegetation at Overstrand blowing in the wind), taking care to capture high quality imagery may not improve the results of the survey, and so limited focus should be given to the image quality in these circumstances.

Chapter 6: Greyscaling

This chapter details the results of varying a single greyscale channel within image registration and bundle adjustment, all the way through to dense cloud production using the standard workflow described in the methods (Figure 3.9). This chapter is split up into three sections, one focussing on quantitative results between clouds at the same sites and how error and point counts vary depending upon the greyscale channel used. Secondly, a more qualitative approach is taken, as clouds are investigated spatially to investigate areas within the clouds which have not been faithfully represented, and to investigate the differences between them in this circumstance. Lastly, the results from the non-linear image processing algorithms are presented individually, compared with the original RGB image set.

6.1 Image inputs

The intention of these tests is to investigate underlying trends which might be present within each of the three colour channels (Figure 6.1) and at a very high level investigate how interactions between these channels and the workflow might be affecting the quality of photogrammetric products.

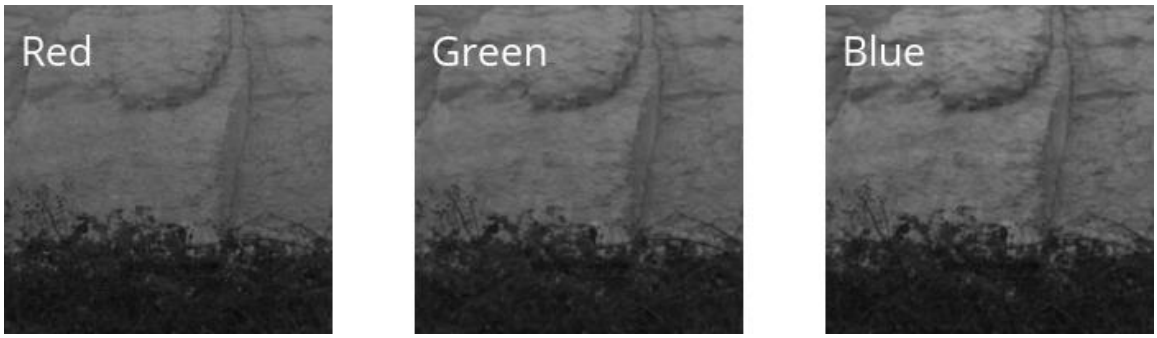

Figure 6.1 Primary colour channels within an RGB image. Virtually no difference can be perceived by the human eye.

As discussed within the literature review, single channel images can be produced using algorithmically produced images starting from three band RGB images (Figure 6.2).

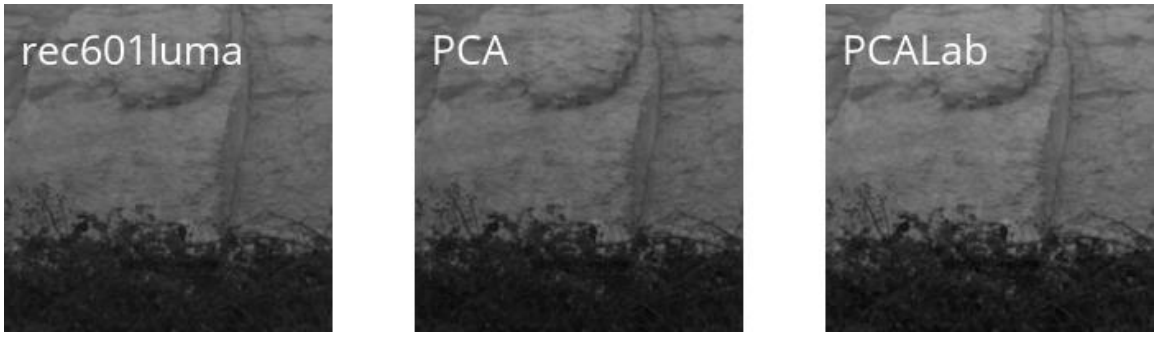

Figure 6.2 The rec601 luma (as described in Section 2.3.4) is a commonly used RGB-grey transformation. It and the two non-linear operators used show, as with the RGB channels, virtually no difference.

6.2 Summary of statistics

For the greyscale tests for the Hunstanton survey, the range of values for the GCP error was from 5.29 mm to 5.54 mm (Table 6.1). The highest GCP error was 5.54 mm for the green image set, with the lowest being 5.29 mm for the PCA image set. The RMSE-P ranged from 0.25 pixels to 0.36 pixels, the highest being from the red set, 44% higher than the lowest, the PCALab set. PCALab returning the lowest RMSE-P was consistend that results which has been reported in Verhoeven et al., (2015). The highest median error was found in the red set at 6.46 mm, just a 3.8% difference between it and the highest, the RGB set, with an error of 6.22 mm. Lastly, point count ranged from 18.24 million to 19.8 million, with the highest count being in the RGB image block, the lowest in the red image block.

These metrics suggest only extremely small differences exist between image sets depending on the way the single greyscale channel is generated. While certain techniques, such as the PCALab image set, show a reduction in residual error in image registration, this does not appear to translate into either an improvement in the median error or the point count.

For the Overstrand survey, the range of values for the GCP error was from 8.17 mm to 9.27 mm. The highest GCP error was 9.27 mm for the PCA image set, with the lowest being 8.17 mm for the red image set. The highest median error was found in the red set at 21.06 mm, a 15 % difference between it and the highest in the PCALab set with an error of 18.33 mm. The RMSE-P ranged from 0.379 pixels to 0.481 pixels, the highest being from the red set, 27 % higher than the lowest, the PCALab set. Lastly, point count ranged from 18.94 million to 21.44 million, with the highest count being in the RGB image block, the lowest in the red image block.

The improvement in RMSE-P which was present within the Hunstanton experiments is not evident within the Overstrand results, which show a reasonably consistent RMSE-P between greyscale

treatments. The PCALab image set does return a lower median error than each other image block in the group of experiments, though at the expense of the point count which is lower than that of the RGB images.

Table 6.1 Summary of results from greyscaling tests

Site	Set	Mean GCP error (mm)	RMSE-P	Median error (mm)	Point count (millions)
Hunstanton	Red	5.34	0.36	6.46	18.24
Hunstanton	Green	5.54	0.26	6.34	19.22
Hunstanton	Blue	5.33	0.30	6.23	18.69
Hunstanton	PCA	5.29	0.26	6.32	19.28
Hunstanton	PCALab	5.38	0.25	6.27	19.40
Hunstanton	RGB	5.36	0.28	6.22	19.80
Overstrand	Red	8.17	0.48	21.06	18.94
Overstrand	Green	8.88	0.40	20.87	19.61
Overstrand	Blue	9.25	0.44	20.25	19.66
Overstrand	PCA	9.27	0.41	19.11	19.43
Overstrand	PCALab	9.24	0.38	19.43	20.15
Overstrand	RGB	8.38	0.39	20.20	21.44

6.3 Colour differences

The results from the linear combinations of RGB channels making up a greyscale image are presented using a novel ternary diagram, where the coefficients can be traced to each axis to keep track of each RGB coefficient making up that greyscale channel. These diagrams show differences between the image sets in a triangular format. The area represents various weighted greyscales made up of linear combinations of the three bands in the form:

$$image = \mathrm{x} * \mathrm{R} + \mathrm{y} * \mathrm{G} + \mathrm{z} * \mathrm{B} \quad [6]$$

Where x, y and z are weights associated with a given derived image which must sum to 1. These weights were changed in 0.1 steps, covering the whole subspace.

6.3.1 Median Error

6.3.1.1 Hunstanton

The error seen within the cloud to cloud differences at Hunstanton ranged from 6.18 mm to 6.52 mm (Figure 6.3). Between clouds little variation is seen, though the minimum occurs around the blue channel. Considering the nature of the Bayer array (Figure 2.4), this is counter-intuitive, as the effective sampling rate is expected to be twice as high in the green band when compared with either the blue or red channel, and with this higher sampling rate error would be expected to be lower.

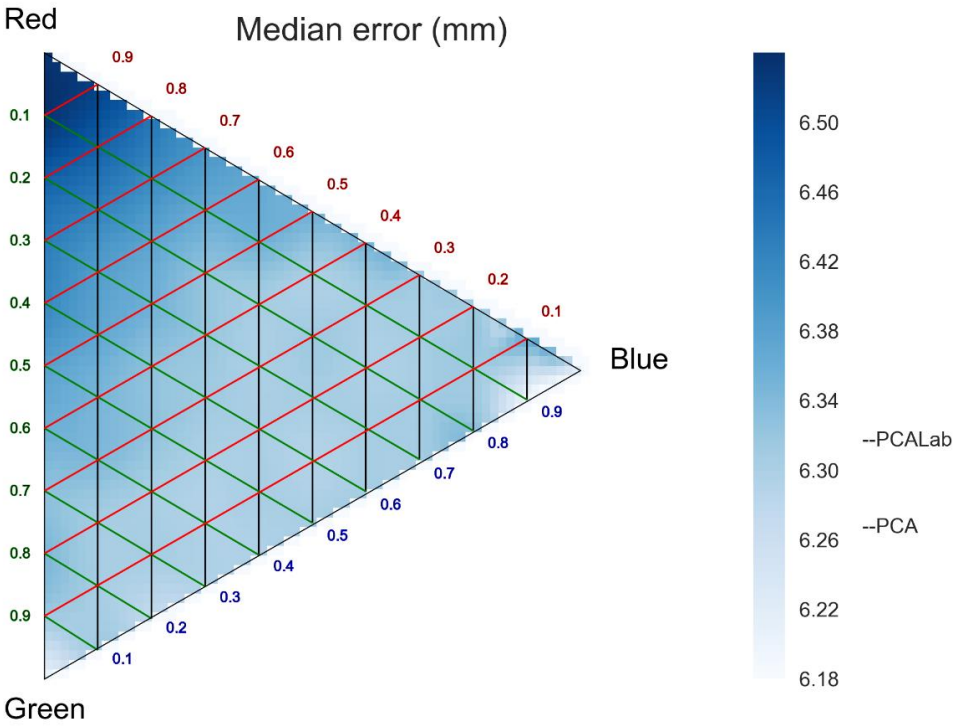

Figure 6.3 Median error for combinations of the individual RGB channels for the Hunstanton survey shows gradients favouring Blue/Green combinations of greyscale channels.

6.3.1.2 Overstrand

Within the Overstrand experiments the results are more heterogeneous, with higher levels of error being seen in the pure green and blue channels when compared with the Hunstanton results (Figure 6.4). The range of error is between 18.74 mm to 20.91 mm across all the image subsets sampled, with large discrepancies seen between extremely similar (in terms of band combination) image blocks. This is likely a result of stochastic effects because of the non-linear nature of the solver used within the SfM workflow (as discussed in Chapter 4). Speaking generally, minima seem to be concentrated in areas where there are large amounts of mixing of the bands. At each of the red, green and blue extremes we see large errors.

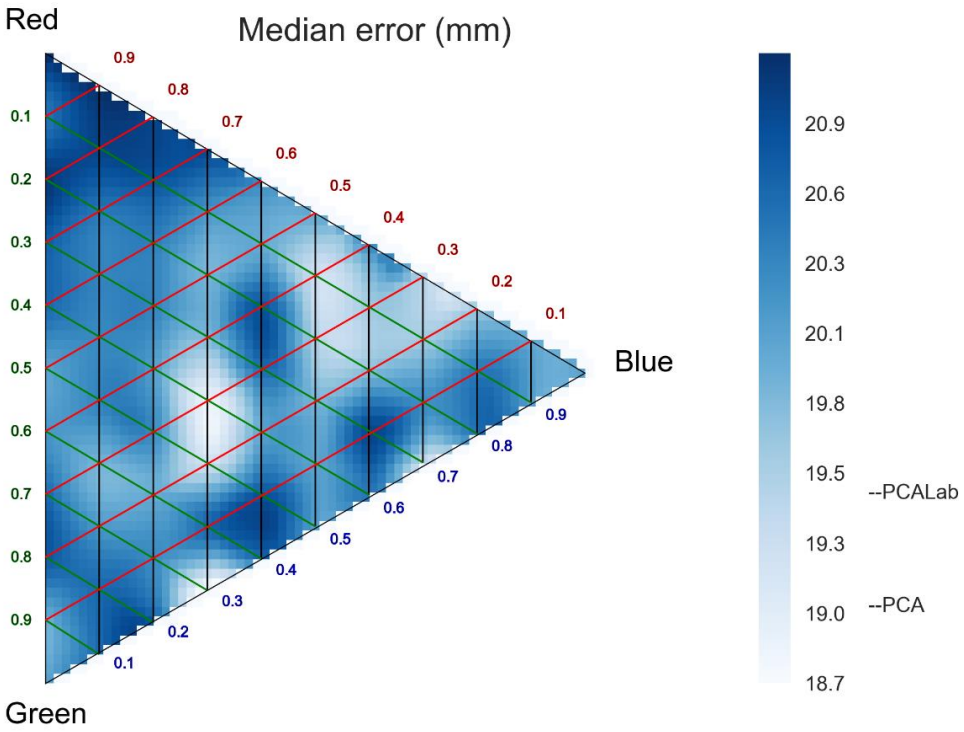

Figure 6.4 Median error for combinations of the individual RGB channels for the Overstrand survey don't show any gradients, with minima found, generally, in areas where the colour channels are mixed.

6.3.2 Density

6.3.2.1 Hunstanton

Point density for the Hunstanton results has a very narrow range, from 714 to 723 points, representing a difference of just 1.2% between the densest and sparsest point cloud across all the sets analysed (Figure 6.5). In general, a trend in the point clouds becoming thinner towards the red channel exists, with density tending to increase as more of the blue channel is added. The green channel does not show the level of density expected when compared with each of the other primary colour channels, and the areas around the pure green channel on the error surface show a sparsity comparable with that of the red channel. Greyscale channels with larger amounts of the blue channel in the linear combination perform well when compared with the other two channels.

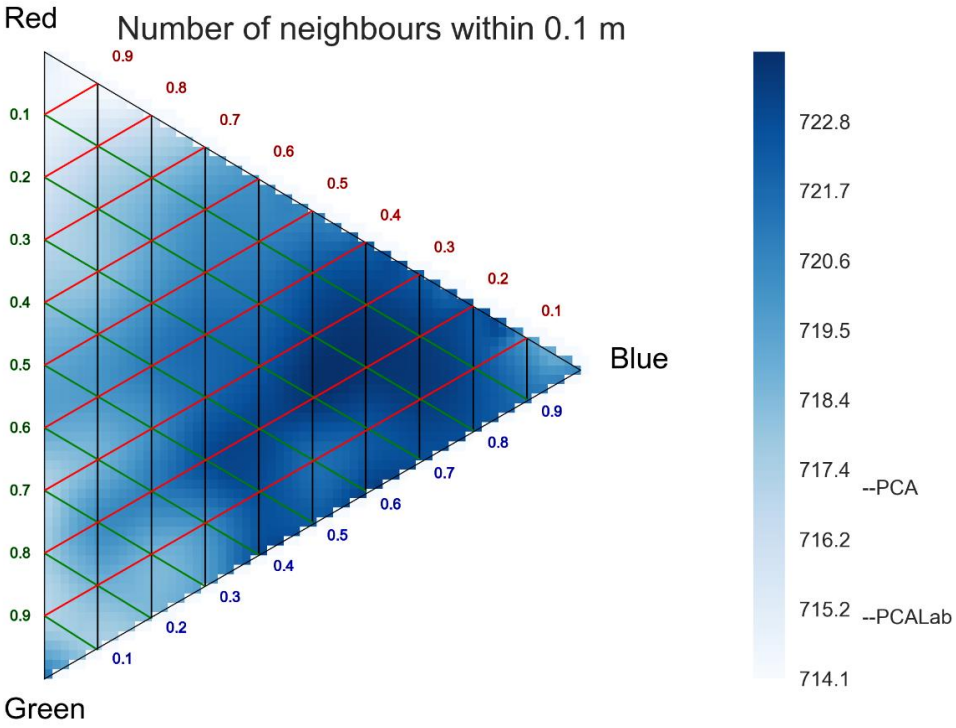

Figure 6.5 Number of neighbours for combinations of the individual RGB channels for the Hunstanton survey. Blue weighted channels show a higher density, suggesting they are more highly clustered than the other channels.

6.3.2.2 Overstrand

For the density of the Overstrand single channel sets, comparable results to that of Hunstanton over parts of the density surface were seen (Figure 6.6). The range of density values is higher than that of Hunstanton, with the lowest being 210.3 and the highest 221.35, representing a difference of 5.25%, much larger than that of the 1.2% seen at Hunstanton. As with Hunstanton, the red channel tends to produce sparser point clouds, with a concentration of sparser clouds shown on the plot in the area surrounding the pure red channel. Again, as with Hunstanton, the blue channel tends to produce denser clouds than the red channel, though at Overstrand the samples including large amounts of the green channel by far produced the densest clouds, with a gradient being evident within the values retrieved.

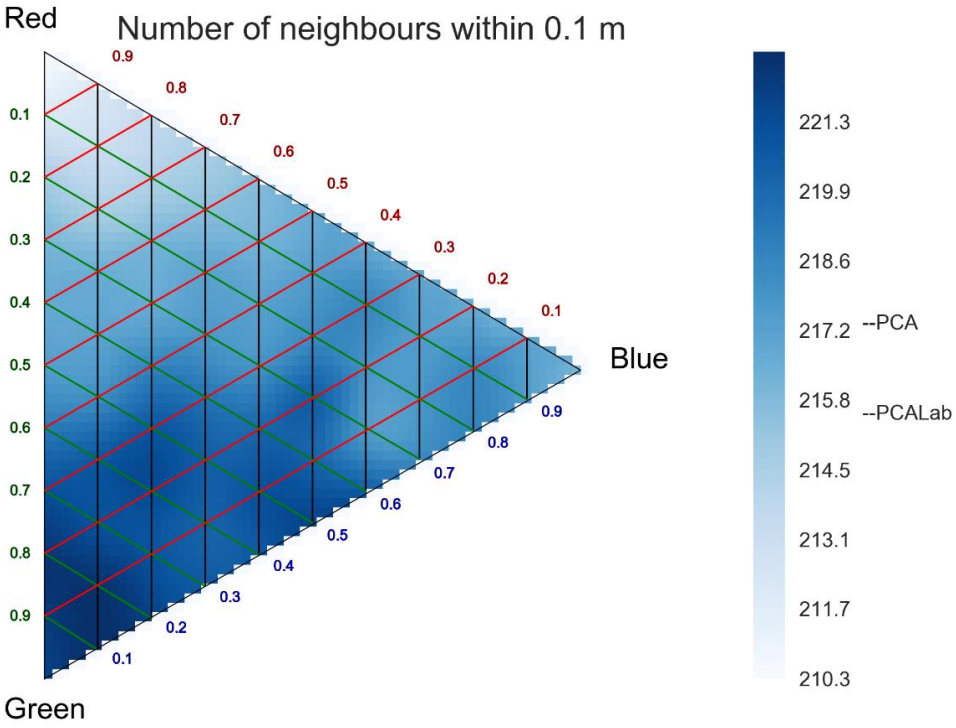

Figure 6.6 Number of neighbours for combinations of the individual RGB channels for the Overstrand survey, a trend is seen for green weighted channels to produce the densest clouds.

6.3.3 Point count

6.3.3.1 Hunstanton

Intuitively, point count and point density are highly correlated, though point counts show a slightly higher range in percentage terms (Figure 6.7). The minimum point count for the Hunstanton tests within the experiments presented is 18.22 million for the pure red channel to 19.34 million for a combination including 80% of the green channel and 10 % of each other channel. Compared with density, point counts are generally higher in the band combinations including more of the green channel, which wasn't the case with the density metrics presented previously. Thus, denser clouds with fewer points in them are being produced by channels from majority-blue band combinations, and thinner clouds with more points by channels from majority-green band combinations at Hunstanton. This indicates the points in majority-green band combinations are spread out more widely, which is the optimal situation, whereas this is not the case for the majority-blue band combinations.

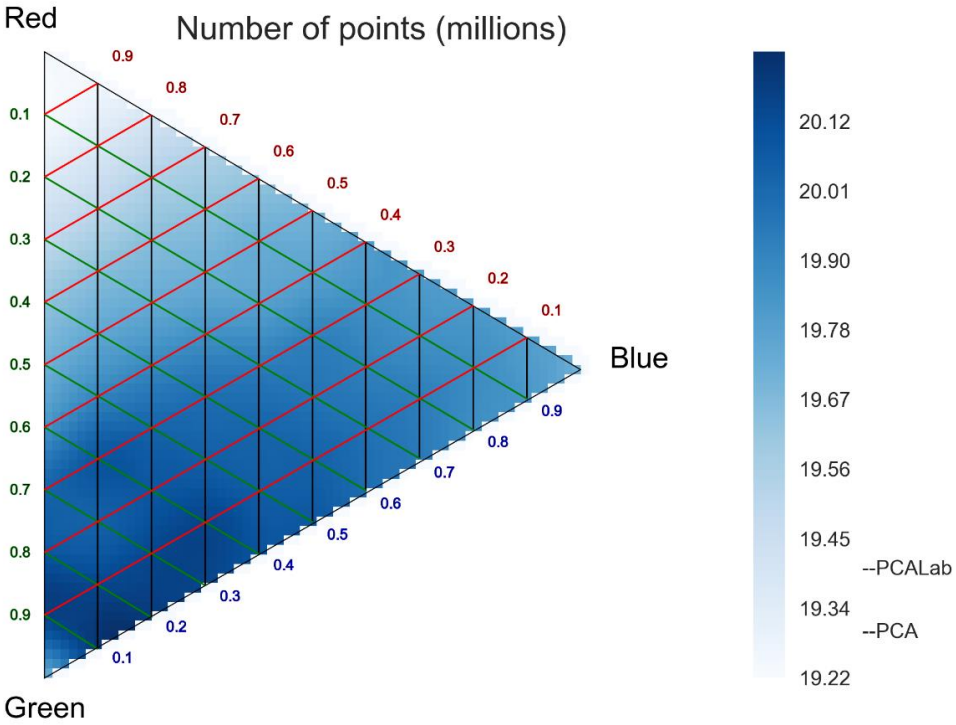

Figure 6.7 Number of points for combinations of the individual RGB channels for the Hunstanton survey, green-weighted channels produce the highest number of points.

6.3.3.2 Overstrand

With the Overstrand results, point counts are seen to be very well correlated with the Hunstanton survey results (Figure 6.8). The range of point counts is wider at Overstrand when compared with Hunstanton, with a minimum of 17.94 million points in the red channel to 19.61 million points for the green channel, representing a difference of 9.3 %, much larger than the difference in density range seen. As with Hunstanton, higher point counts were associated with combinations including large amounts of the green channel.

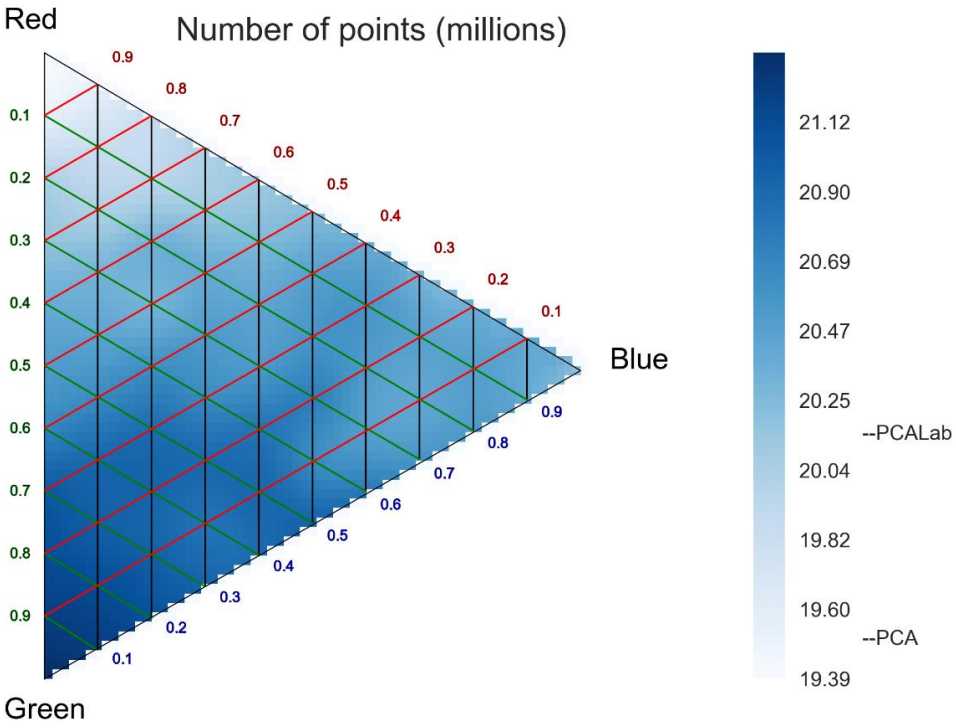

Figure 6.8 Number of neighbours for combinations of the individual RGB channels for the Overstrand survey, green weighted channels produce the highest number of points.

6.4 Lens model variation

As discussed in the Section 3.4.5, within bundle adjustment a lens model is fit to enable image orthorectification, as well as allow correction of focal length (as compared with the EXIF file) and location of the principal point of the sensor. Within the research design, a separate floating lens model was fitted for each photogrammetric block.

6.4.1 Hunstanton

The differences in focal lengths for each image block were visualized to ascertain any effect that the image processing may be having on how the camera lens is optimized. From this, we can see a very ordered shift between each of the colour bands, with the focal length of each band being located approximately 0.3 mm from one another (Figure 6.9). Interestingly, the trend seems to account for any potential chromatic aberration, the differences between images formed in each colour band because of slight differences in the nodal point of the lens. The linear trend from blue to red across the central point would be expected as the wavelength of the light increases, and so the focal length due to chromatic aberration. The optimizer accounts for this discrepancy by setting a longer focal length in the camera model for red weighted bands. The difference in the X and Y axes shows the software attempting to account for a perceived difference in focal lengths for each axis, this asymmetry can often be caused by defects in lens design.

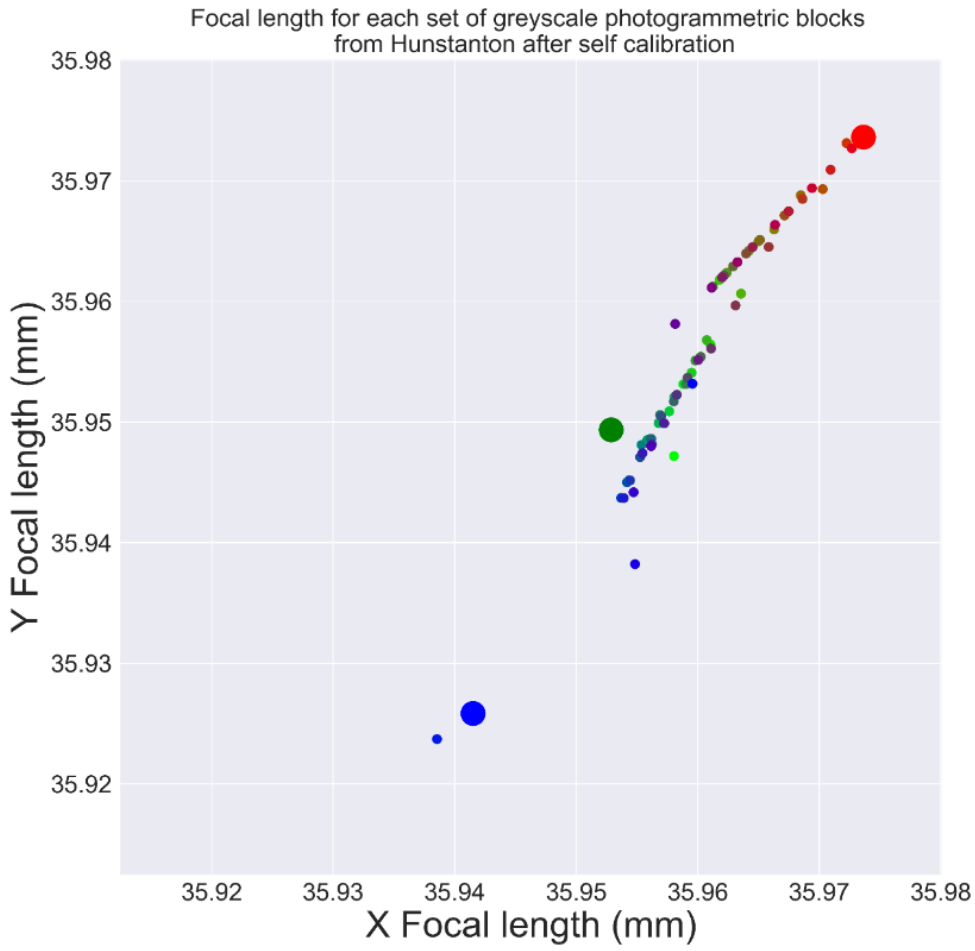

Figure 6.9 Focal lengths for lens models produced from blocks of images using single channels made up of combinations of the individual RGB channels for the Hunstanton survey. Red, green and blue individual channels are denoted by the large circles, with each colour the RGB colour combination making up that image set

6.4.2 Overstrand

For the Overstrand experiments the same trend is seen, though the asymmetry reported for the Hunstanton results are not seen owing to the camera position/model being set before image matching, and so the focal length is the same in both X and Y (Figure 6.10). The effect of chromatic aberration appears to be reduced, as the red, green and blue bands show less of a spread than that seen for the Hunstanton results. The range of focal lengths reported by the software's optimizer is comparable for both sites, though for the Hunstanton results the focal length is generally longer (by approximately 0.1 mm). Certain solutions of band combinations appearing as apparent outliers is likely a result of the stochastic nature of the non-linear optimizer and are not accounted for directly within this experimental design.

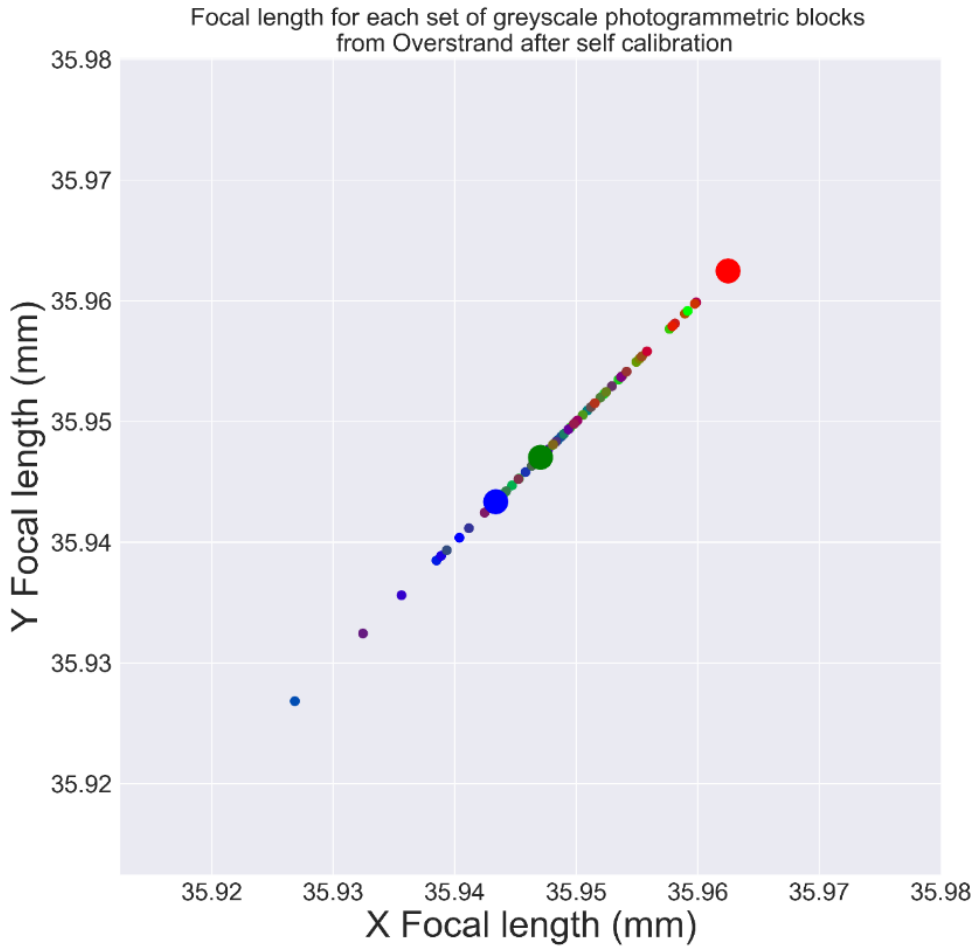

Figure 6.10 Focal lengths for lens models produced from blocks of images using single channels made up of combinations of the individual RGB channels for the Overstrand survey. Red green and blue individual channels are denoted by the large circles, with each colour the RGB colour combination making up that image set

6.5 Spatial analysis

6.5.1 Red, green and blue channels

Differences between the primary colour channels are presented to analyse whether any spatial trends can be identified between each of the red, green and blue image sets. Figure 6.11 shows a crop Section from Hunstanton, highlighting a well textured area of the upper cliff, with both level of error and point density presented for each of the channels. In terms of cell occupancy in this area, we see almost identical gaps within the DSMs produced. Error is concentrated around the upper left portion of the crop section, with error in the set containing the blue channel images showing a slightly better fit to the truth when compared with the other two. In Figure 6.12, towards the bottom of the dark carstone where differences might be expected due to differing colour contrast, we see almost identical produces from each of the three colour bands in both metrics presented.

For the Overstrand DSMs, a textured area towards the back of the landslide is highlighted in crop sections, which shows, again, each of the three-individual red, green and blue bands. Only very slight differences exist between each of the three channels, though some thinning appears to exist in the green channel, where pixel occupancy is reduced towards the upper left portion of this area (Figure 6.11). For the areas towards the right of the crop section, point density is slightly higher in the photogrammetric set made up of the green image channel. In Figure 6.12 towards the front of the landslide we see nearly identical products in all colour channels, reinforcing the idea that differences between products generated from custom greyscale channels are very small.

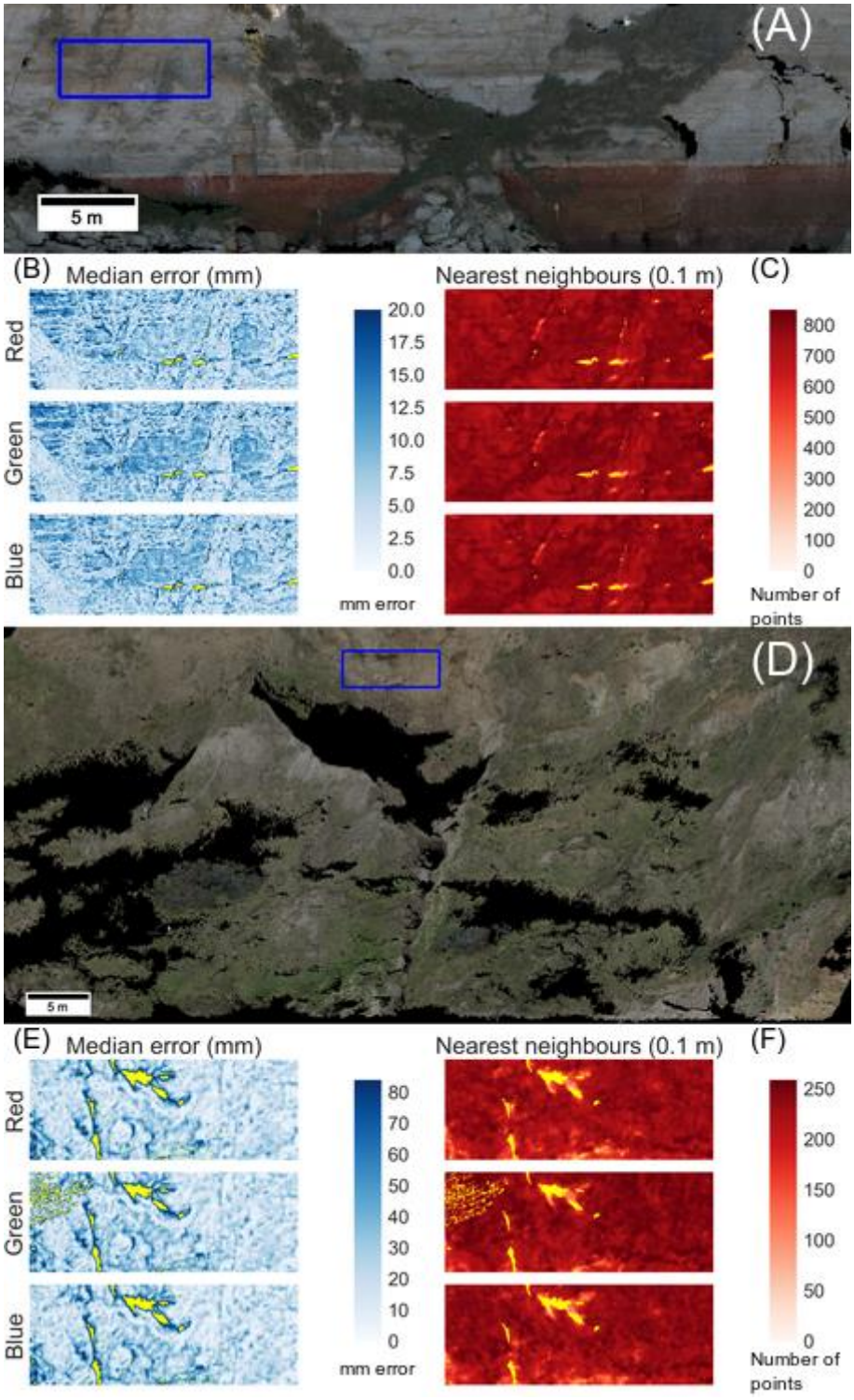

Figure 6.11 Differences from point clouds produced by the three primary colour channels are only very slightly for the Hunstanton results. For Overstrand some unoccupied pixels are noted in the reconstruction of the back wall (E). No data (missing) values are in yellow.

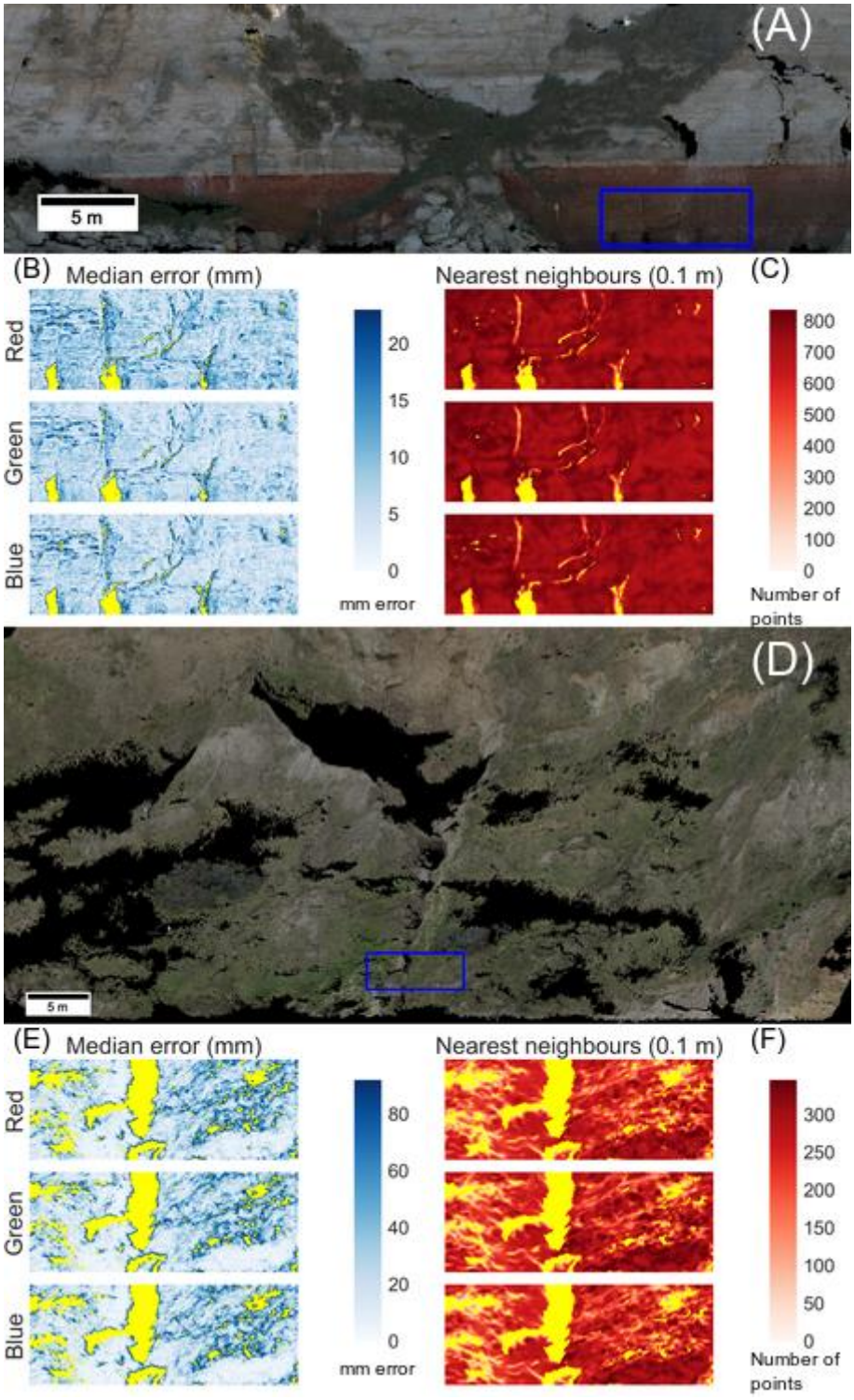

Figure 6.12 Differences from point clouds produced by the three primary colour channels are only very slightly different for both sites. No data (missing) values are in yellow.

6.5.2 RGB, PCA and PCALab

Differences between the untreated RGB input and both PCA based greyscale channel generation algorithms are presented in Figures 6.13 and 6.14. In the upper portions of the Hunstanton cliff, both error and density are almost identical between each of the three treatments presented, as the metrics in Table 6.1 would suggest. Investigating the region of carstone towards the lower right, the PCALab image set holes begin to open up in an unordered fashion, suggesting the texture of this region is not being adequately recovered in the derived images (Figure 6.14 B).

For the Overstrand results, as with the green channel previously, unoccupied cells within the DEM are seen towards the upper regions of the landslide which are not present in either of the PCA based image sets, though were not in either of the red or blue channel DSMs either. This suggests that this effect is a result of differences within the image registration and bundle adjustment stages which may not be specific to the greyscaling operation. At the front of the landslide almost identical products are again seen.

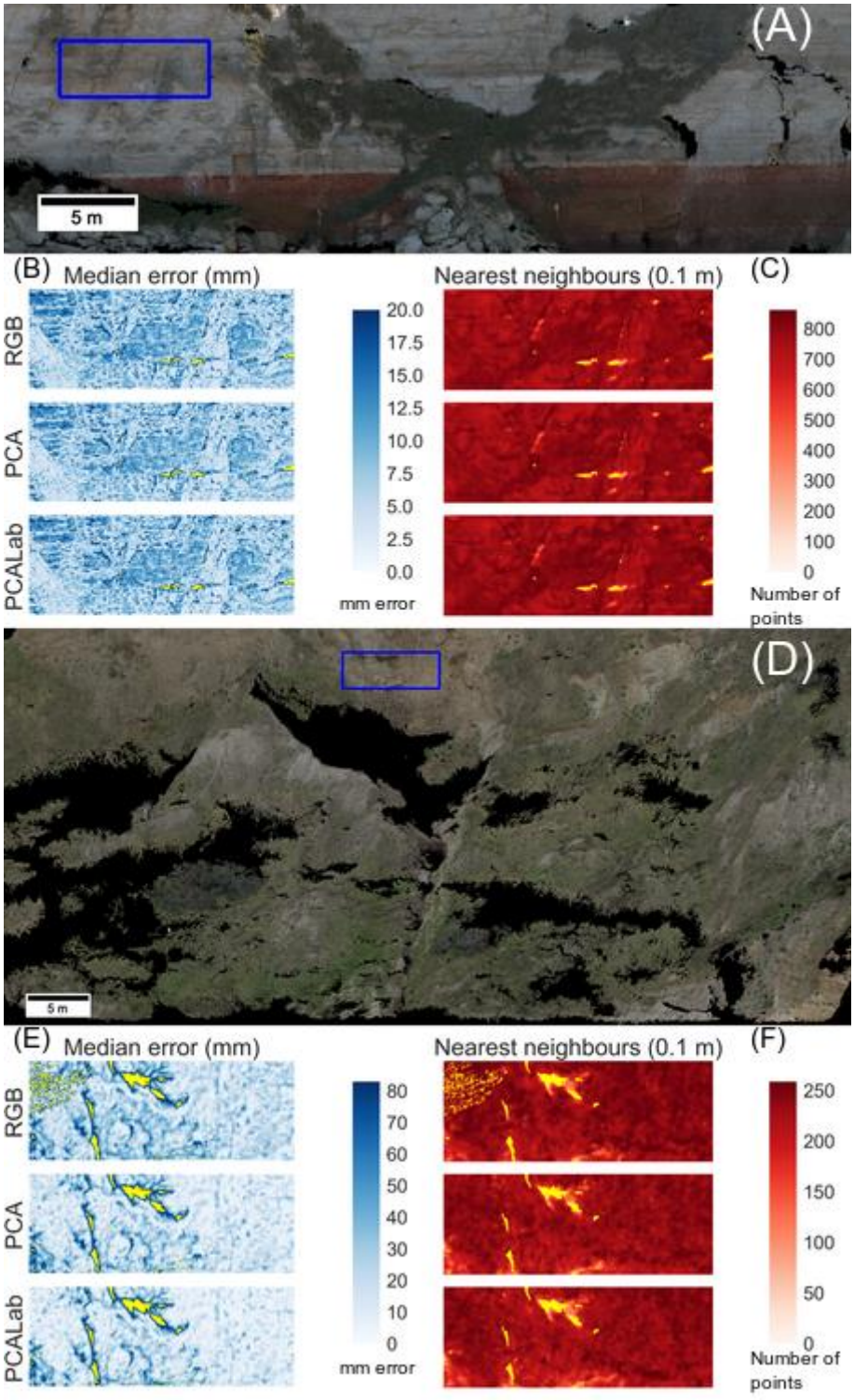

Figure 6.13 The two PCA-based algorithms are extremely similar to the untreated set for the Hunstanton results (A), with slight variation for the back wall of Overstrand in terms of pixel occupancy (E), a similar trend to this was seen for the individual Green channel's results (Figure 5.11 (E)). No data (missing) values are in yellow.

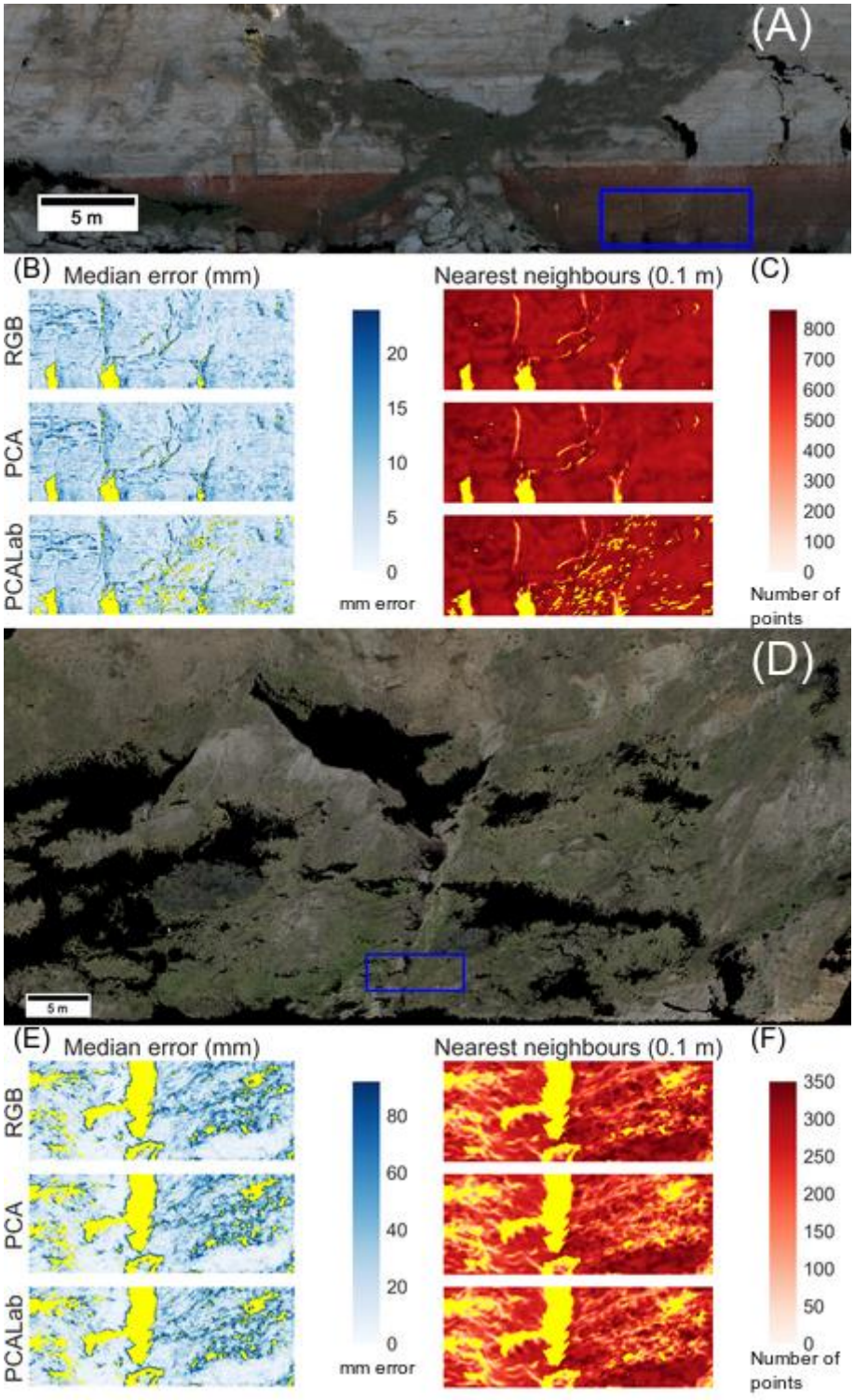

Figure 6.14 PCALab underperformed in areas of dark carstone at Hunstanton, which, when combined with Figure 6.13, suggest the PCA set was the most consistent. All of the image sets performed relatively well when it came to reconstruction of the gully towards the front of the landslide. No data (missing) values are in yellow.

6.6 Distribution analysis

Results showing the histograms of errors and nearest neighbours reveal very little variation between point clouds produced from each greyscale image set when compared with the unaltered RGB images (Figure 6.15). At Hunstanton, the error quartiles almost perfectly coincide across all the image sets, with the PCALab set showing a slightly higher density than any of the other sets, though this could be due to outliers not being included within the final product.

At Overstrand, the PCALab set produces a slightly better cloud in terms of error, though this again is likely due to outlier exclusion within the dense matching algorithm. In general, the sets produce very similar results, suggesting any improvement in the quality of products at a site such as Overstrand would be very minimal.

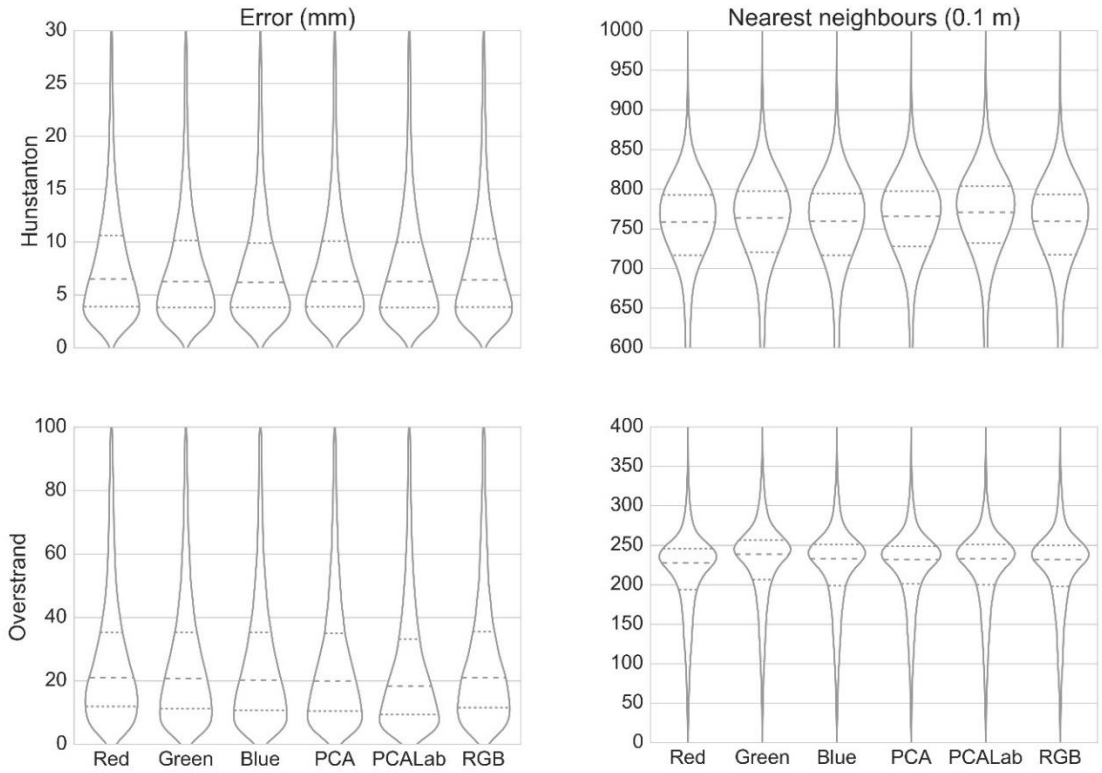

Figure 6.15 Clouds produced from images varying the band used for image creation show very little difference, though the red channel perform slightly worse than the others.

6.7 Difference of DSMs

Difference of DSMs, as in the image quality chapter, were performed on the best and worst performing image sets within the space of linear combinations (Figure 6.16). At Hunstanton, this was a single channel including a large proportion of the blue channel performing best, with the red-weighted channels performing relatively poorly, in terms of median error. For the Hunstanton results, unlike the image quality results, some spatial trends appear. The blue weighted channel generally performs better in dark areas, such as the carstone region to the bottom right of the DSMs, and the bare rock face to the left of the image block. Conversely, the red-weight channel appears to perform much better reconstructing the rock face towards the back wall above the landslip.

For the Overstrand results no spatial trend is seen, with the rock face towards the front of the landslide showing the greatest difference (Figure 6.17). The worst performing sets in terms of error were the pure red channel, with the 0.3 red, 0.4 green and 0.3 blue channel performing best in this regard.

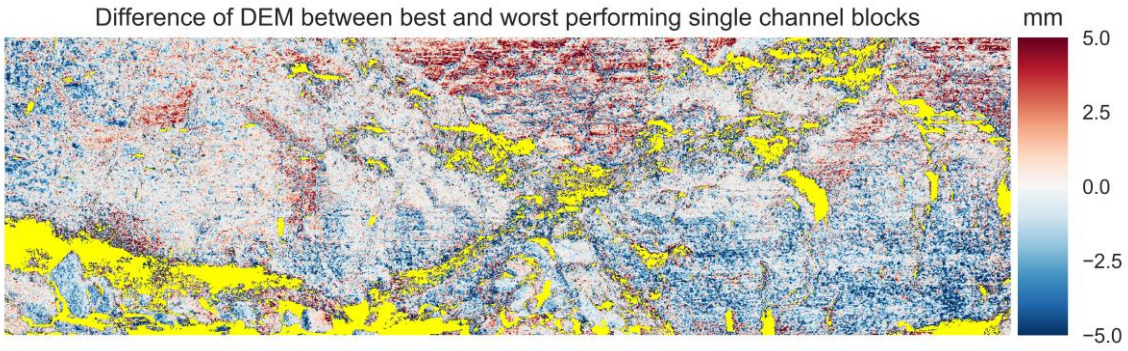

Figure 6.16 A spatial trend is seen between the blue and red weighted channels for Hunstanton. Most of the difference appears in the Carstone at the bottom right, though differences are, in general, only very slight. No data (missing) values are in yellow.

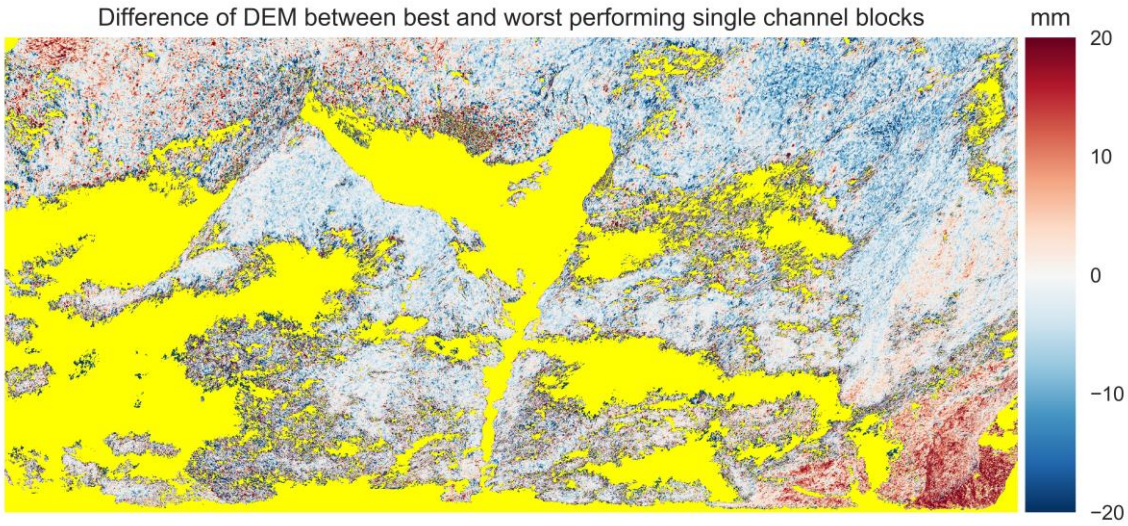

Figure 6.17 No obvious spatial trend is seen with the best and worst clouds for Overstrand. No data (missing) values are in yellow.

6.8 Summary

The single channel experiments suggest very little improvement can be made through image pre-processing of individual images, when compared with the original, uncompressed RGB image set. While some slight spatial variations were seen in each site, the high-level statistics allude to this similarity.

Chapter 7: Geometry

7.1 Introduction

This chapter details the results of varying combinations of images taken from a range of different perspectives for the two study sites detailed using the standard workflow described in the methods (Figure 3.4). The aim of this chapter is to establish whether viewing geometries had a measurable impact on the quality of the derived products from both sites.

7.2 Image inputs

Images acquired from both sites were split into image sets depending on the viewing geometry relative to the surface being imaged. Considering the images acquired at each site were collected in a similar manner, on a transect with 5 angles at each of 8 camera stations (labelled 1-8), image sets were generated systematically from the full image block (Table 7.1).

Table 7.1 Summary of image subsets included within the network design tests

Set	Description of image input	Total images
1	Image from all stations from all angles	40
2	Images from all stations, at +30°, nadir and -30° from surface face	24
3	Images from all stations at +15° and -15° from surface face	16
4	Images from stations 1, 3, 5 and 7 from all angles	20
5	Images from stations 1, 3, 5 and 7 at +30°, nadir and -30° from surface face	12
6	Images from stations 1, 3, 5 and 7 at +15° and -15° from surface face	8
7	Images from stations 1, 4 and 7 from all angles	15
8	Images from stations 1, 4 and 7 at +30°, nadir and -30° from surface face	9
9	Images from stations 1, 4 and 7 at +15° and -15° from surface face	6

The objective of the image subsets was to establish any effects on both density and accuracy of the point cloud products resulting from differences within the image networks. The tests vary both the total number of image inputs and the angles at which data were acquired. While certain viewing geometries led to differing occlusions between image sets, in general the coverage for each set is as complete as would be practical in an operational environment.

7.3 Photogrammetric results

7.3.1 Summary of statistics

7.3.1.1 Hunstanton

For the geometry tests from the Hunstanton survey, the range of values for the marker error was from 4.48 mm to 6.47 mm (Table 7.2). The highest marker error was 6.47 mm for image set 4, with the lowest being 4.48 mm for image set 7. The RMSE-P ranged from 0.09 pixels to 0.27 pixels, the highest being from set 1, which contained all of the images within the block, three times as high as the lowest, set 6, which contained just 8 images in total but retained a much lower RMSE-P when compared with any other block within the tests. Speaking generally, a trend of RMSE-P being lower in image sets with fewer images was seen. The highest median error was found in set 4 at 7.62 mm, representing a 29 % difference between it and the highest in the set 6 with an error of 5.9 mm. Set 6 therefore returned not only the lowest residual error within image matching but also the lowest median error when compared with every other image block. Lastly, point count ranged from 10.5 million to 19.8 million, with the highest count being in set 1 which contained all images, the lowest in set 8, which contained 50 % more images than set 9 but owing to variations in viewing geometry returned less points in dense matching.

Table 7.2. Summary of results for the geometry tests from the Hunstanton survey

Set	Mean marker error (mm)	RMSE-P	Median error (mm)	Point count (millions)
1	5.35	0.27	6.36	19.8
2	4.97	0.23	6.41	17.1
3	5.77	0.19	6.65	17.6
4	6.47	0.23	7.62	15.2
5	6.16	0.19	5.91	13.4
6	5.32	0.09	5.90	13.1
7	4.48	0.22	6.38	12.8
8	4.93	0.18	5.95	10.5
9	4.93	0.15	5.92	10.6

7.3.1.2 Overstrand

For the Overstrand survey, the range of values for the marker error was from 7.68 mm to 13.0 mm, which was, as with the other tests, higher than at Hunstanton (Table 7.3). The highest marker error was 13.0 mm for set 5, with the lowest being 7.68 mm for set 6. Set 6 also performed well in terms of median error (14.43 mm). The RMSE-P in image matching ranged from 0.215 pixels to 0.386 pixels, the highest being from set 1, 80% higher than the lowest, set 7. The highest median error was found in set 1 at 20.2 mm, a 52% difference between it and the highest in set 9 with an error of 13.29 mm. Lastly, point count ranged from 4.19 million to 20.44 million, with the highest count being in set 1, the lowest in the set 9.

Table 7.3. Summary of results for the geometry tests from the Overstrand survey

Set	Mean marker error (mm)	RMSE-P	Median error (mm)	Point count (millions)
1	8.38	0.39	20.20	20.44
2	9.4	0.34	18.89	16.77
3	8.28	0.29	16.71	13.82
4	10.5	0.34	15.60	10.08
5	13	0.29	16.00	7.96
6	7.68	0.27	14.43	7.08
7	9.125	0.22	18.17	6.97
8	10.2	0.33	20.02	5.16
9	10.5	0.34	13.29	4.19

7.3.1.3 Intersite comparsion

Comparing the two sites, there is a trend of marker error increasing with image sets containing fewer images, though when very few images are present (such as set 9) this appears to fall again for the Hunstanton site. The lowest RMSE-P was seen in set 6 at Hunstanton, which contains only 8 images, registering relatively well due to the convergent viewing geometry along the y axis and large overlap. For the Overstrand results the minimum RMSE-P was seen in set 7, which contained 15 images. No consistent relationship can be drawn between the RMSE-P as the results varied so much between sites, though as with the marker error, though the results tended to be lower in sets containing fewer images. In terms of error, as with the two previously mentioned metrics, there appears to be a reduction in the median error with decreasing amount of input images.

Point count at both sites was highly related to how many input images there were in each set, though this did not always hold true, as sets 7-9 appeared to perform very poorly due to the very wide spacing of the camera stations. Lower median error was seen, in general, in clouds containing fewer points, which suggests that more certain, but fewer, points were being included within the dense clouds generated in these tests.

7.3.2 Spatial variability

7.3.2.1 Hunstanton

For the Hunstanton results, large discrepancies were seen in two areas of the cliff (Figures 7.1 and 7.2). In the upper portions towards the top left of the image, further crop sections are presented showing a similar trend, though in this area where there is lots of texture, the acute differences seen in the lower portions of the image are not observed. The differences in occupancy between sets 1 and 6 is very minor, with a few small unoccupied regions from set 1 becoming bigger in set 6, and much more noticeably so in set 9, which contains only 6 images. Point densities show the expected trend of set 1 having a very consistent higher density when compared with both the other sets presented. The differences between sets 6 and 9 are significant, suggesting the viewing positions of set 6 was one of the bigger factors in determining the coverage of this Section of the cliff.

Within the region to the lower right of the cliff an area of dark red carstone shows a very acute thinning out in terms of point count (Figure 7.2 C) for sets 6 and 9. The number of images containing the section is reduced for the sets with fewer images, but large holes also begin to open up in the images sets containing fewer images. This thinning could be associated with the general underexposure and lack of texture associated with the region, which increases image registration residuals and subsequently will impact the accuracy of the dense matching algorithm. Areas with higher error in set 1, which contains the most images in the photogrammetric block and has the highest point count, tend to not be occupied in the sets with fewer images. There is a thinning of the point cloud towards the outer edges of the region (towards the right of the crop sections) where image overlap is reduced, and point density is subsequently reduced also. Areas of higher depth discontinuity, such as areas where rockfall has recently occurred, show a more faithful representation in set 1, and in these regions the sets with fewer images contain significant numbers of unoccupied cells.

7.3.2.2 Overstrand

The DSMs produced for the Overstrand results show much larger discrepancies in cell occupancy between the sets when compared with the Hunstanton experiments. Within Figure 7.1, we see

that in a highly textured area towards the back of the landslide, the differences are less apparent, though occupancy is higher in set 1 than the other 2 presented. Set 6, where it didn't return any data from the front corner portion of the landslide, returns a very similar result to set 1, albeit with a lower density for the returned points. This suggests that the images being used to reconstruct this specific area of the landslide are the same in both sets 1 and 6. Within the same region we see set 9, which performed better than set 6 towards the front of the landslide, thinning out relative to the other two sets presented, with gaps in the cloud becoming much more evident.

Looking at Figure 7.2, sets 6 and 9 do not produce adequate data over the middle part of the gully in the centre of the scene. Set 1 produces a much higher occupancy than either of the other 2, and the cell density is also much higher than the other 2, suggesting it is a more appropriate strategy for this more challenging site.

7.3.2.3 Intersite comparsion

The spatial variability between sites reveal that more acute differences are seen in areas where reconstruction is expected to be difficult, within heavily vegetated regions and areas where image overlap is poor. In areas where texture and exposure is adequate, image count doesn't seem to have a major influence on the accuracy of derived products, though density is notably lower. For the Hunstanton results, the surface has more consistent texture and exposure is generally good, reducing image count does still provide reasonable reconstruction, which may interest geoscientists where resources are limited or a time series with a set number of cameras is within a research design. For Overstrand, where texture is lower, the vegetation is heavy and the exposure is generally slightly underexposed, we see the image count and geometry significantly influence the ability of the photogrammetric software to provide an adequate reconstruction for general use.

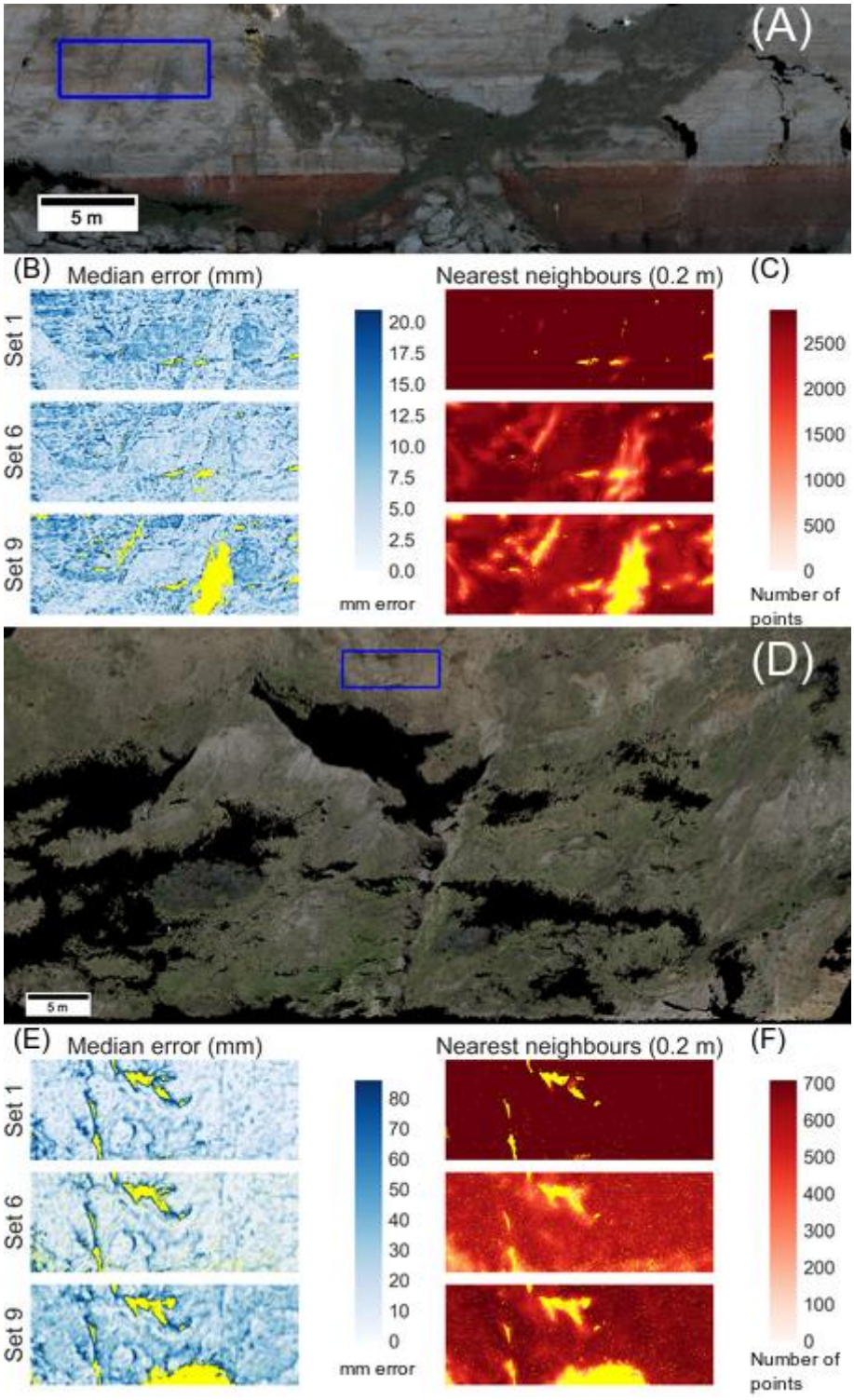

Figure 7.1 Differences between point clouds produced by different imaging configurations. The full set (1) is more complete at both sites, though set 6 shows reasonable performance within both highlighted sections. No data (missing) values are in yellow.

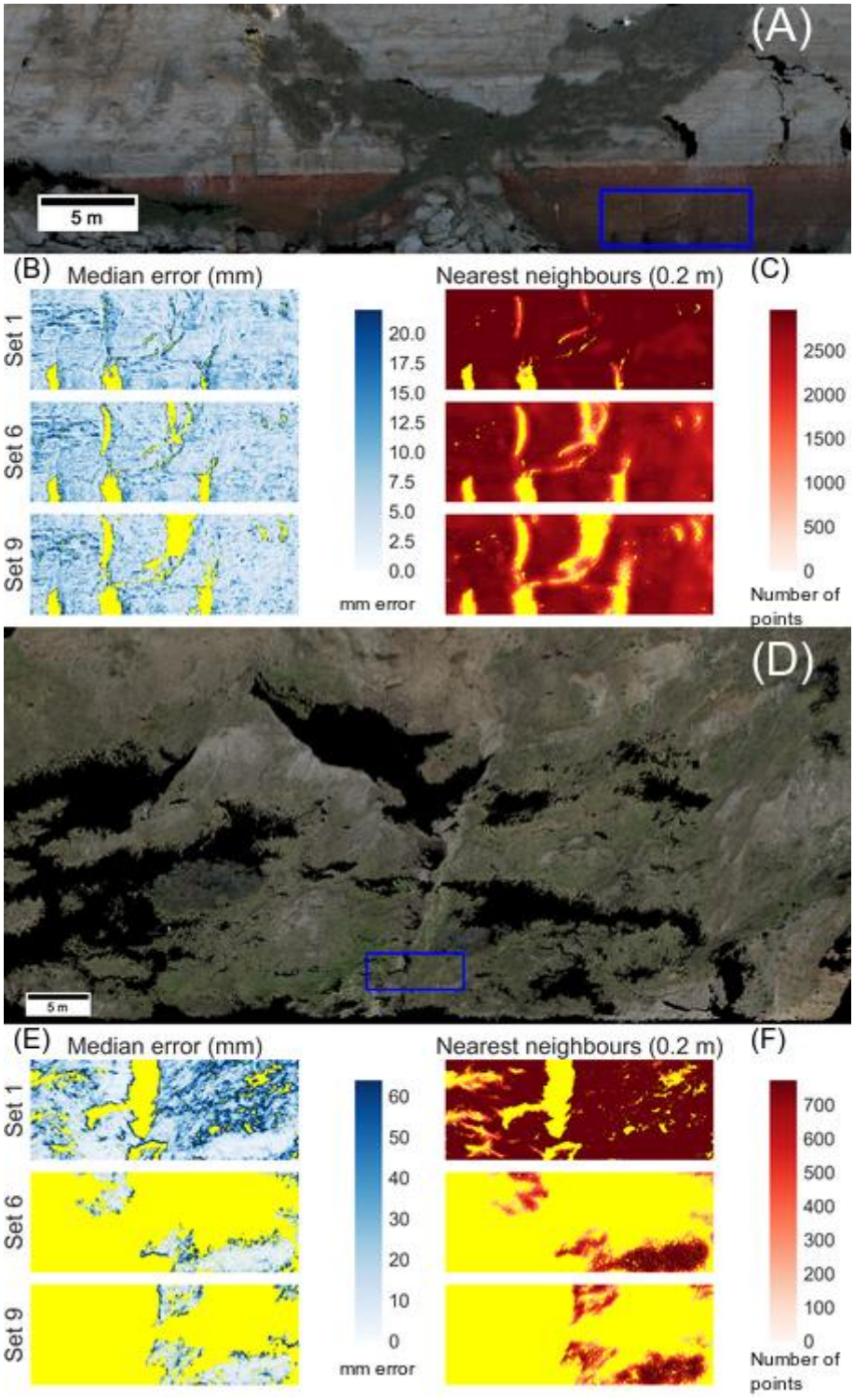

Figure 7.2 Differences between point clouds produced by different imaging configurations. Set 1 is again, the most complete, with sets 6 and 9 completely failing to reconstruct the central middle part where the gully is located. No data (missing) values are in yellow.

7.3.3 Distribution analysis

The distribution of error is close to normal for each set investigated at Hunstanton, though the violin plot reveals that while set 4 contains a relatively large number of images (20), the tail of the plot indicates there are many outliers within the point cloud produced, suggesting that the additional viewing angles degraded the reconstruction quality (Figure 7.3). This is somewhat evident, though less so, in image block 7, which used the same viewing angles as set 4, though from a reduced number of stations. Comparing this with the Overstrand distributions, the same relationship is not seen, with set 4 performing particularly well when compared with the other sets presented.

The density histograms for Hunstanton show a large degree of similarity between certain image sets. Sets 1 and 3, for example, share a very similar shape, though within set 3 the median and quartiles are consistently lower than in set 1, which is to be expected considering set 1 contains 40 images and set 3 just 16. While point count was lower in set 3 (17.6 million) when compared with set 1 (19.8 million), the benefit of both time taken to produce a reconstruction and acquisition time would suggest this strategy appropriate for research where these factors are limiting.

Sets 2 and 4 show a very similar density distribution at Hunstanton while having very different imaging configurations. The image blocks contain similar amounts of images and recovered comparable point counts in the initial reconstruction (17.1 million vs. 15.2 million). However, when error is considered, set 2 outperforms set 4.

The Overstrand density distributions show very different results from that of Hunstanton. While the intuitive trend of blocks containing more images having a higher point density is seen, we do not see the similar density distributions as described in the Hunstanton results. Comparing set 1 with set 3, as with Hunstanton, set 1 shows a much higher median density (973 points within 20 cm on average) when compared with set 3 (765 points within 20 cm on average). Error metrics suggest that while this higher density and point count shows a more complete cloud, the points being added by the extra images are of lower quality, pushing the median error of set 1 higher (20.2 mm) when compared with set 3 (16.71 mm).

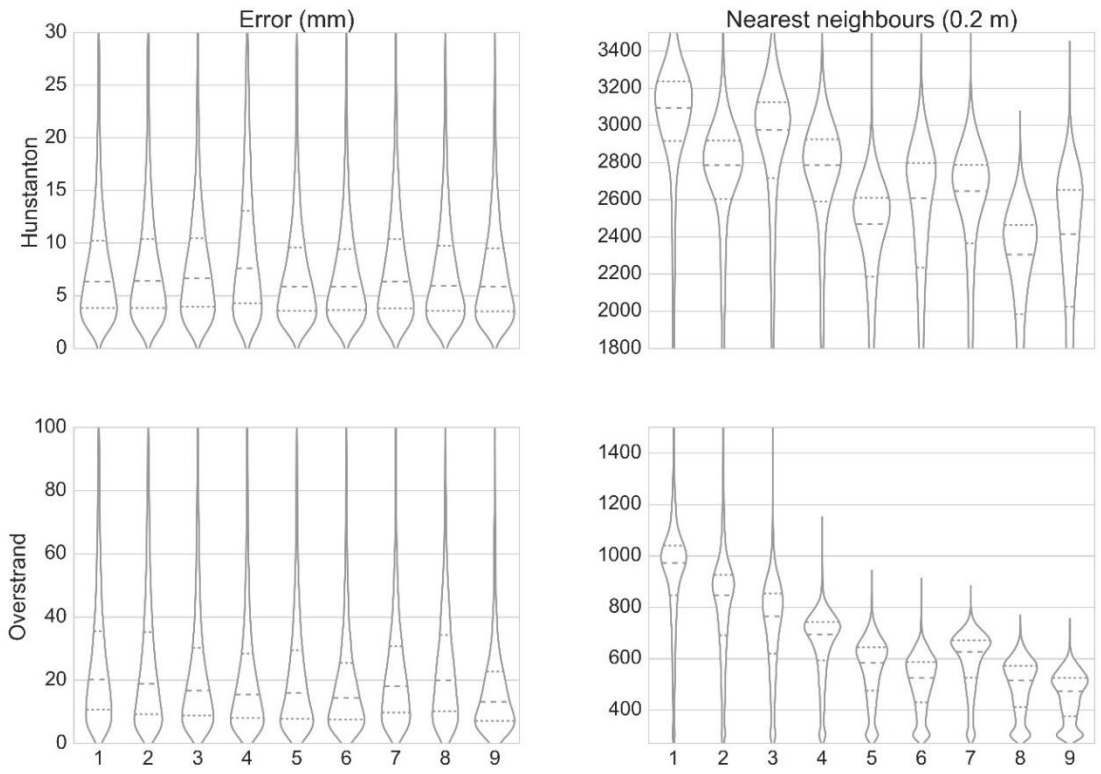

Figure 7.3 Density increases with higher numbers of images in a set, though error is more variable

7.3.4 Difference of DSMs

Considering the emerging trends, sets 6 and 7 were directly compared to ascertain whether the difference in median error was a result of spatial heterogeneity in the areas returned, or whether it performed better across the block. Set 6 performs better along the rock face to the left of the cliff and along the back wall, while set 7 performs better along the landslip itself (Figure 7.4). With this being said, the two clouds are very different in terms of coverage area and appearance as a result of occlusions, so only limited information could be drawn from a direct comparison of error DSMs. For the Overstrand DSMs, no two were deemed worthy for comparison, as large differences in cell occupancy and density made any comparison speculative.

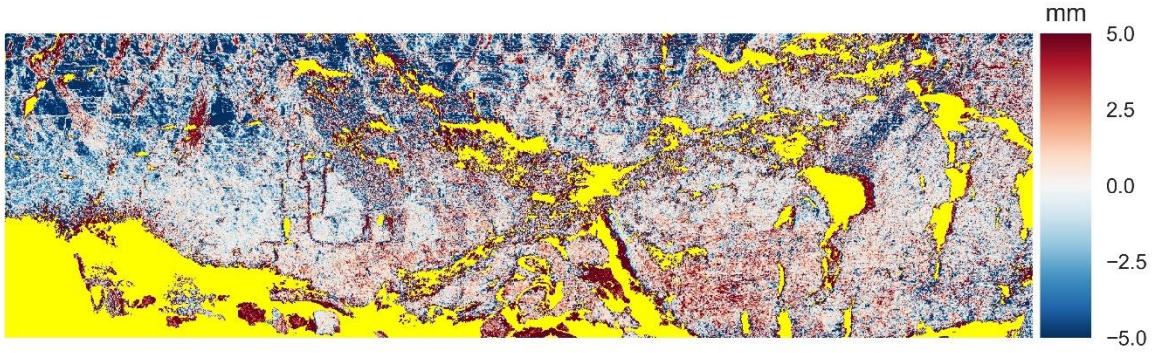

Figure 7.4 Difference of error DEM between sets 6 and 7 for the Hunstanton site. Spatial differences are evident. The upper left side of the cliff face is much more accurately reconstructed within set 6. No data (missing) values are in yellow.

7.4 Summary

The results from the experiments varying network design show that, intuitively, more images tend to lead to more complete and dense clouds, though not necessarily reducing error as might be expected. This is likely a function of the increase in sample size. The additional points likely share a similar distribution (in terms of median error) to those generated from sets with fewer images, and so have little effect on the metric overall. At Hunstanton, where texture across the cliff face was high and exposure adequate, differences between photogrammetric blocks were not very evident. While the set containing the most images did return the densest cloud with highest number of points, other sets with fewer images returned comparable results which would likely be appropriate for different types of research design.

For the Overstrand site, where consistent texture was low and images generally slightly underexposed, the impact of the reduction of number of images was much more apparent with both the point counts and density of the photogrammetric sets.

Chapter 8: Scenario Comparison

8.1 Introduction

The previous chapters have detailed trends within photogrammetric products which correlate with image quality. To generate firm recommendations to the community on the scale of expected degradation at varying image qualities, scenarios were run which parallel what the community would expect to come across depending on the quality of equipment available to them. Initially, each scenario will be described, with a brief introduction to who the user group would be and the level of expertise associated with each. Secondly, images of varying quality are generated depending on the expected level of equipment/expertise associated with each group. Lastly, the results will be described, and noticeable differences highlighted. This exercise seeks to make firm recommendations for how the photogrammetric community should handle data.

8.2 User groups

For each of the photogrammetric products produced within this chapter, the full dataset was used with images being pre-processed using different workflows, according to the respective group. The objective is to highlight how differences which are not necessarily reported within the literature (particularly, for example, the differences between groups 2 and 3 listed here) can influence results.

8.2.1 Amateur users with little experience of image acquisition (Group 1)

Within geoscientific studies, some publications do not detail image acquisition strategies or image quality. As a result, image format, levels of compression or camera settings are unknown (see appendix of O'Connor et al., (2017) for examples of research missing details of image acquisition). Users new to photogrammetry and those not trained in its use would fall under this general description, as would those scoping photogrammetry as a means for primary data collection.

For this scenario, images were reduced to 8 MPs from the native 12-MP resolution, noise of 21 dB was introduced into the image set and the images JPEG-compressed to a 'quality' of 50, meaning that while posterization was still not apparent, significant compression had taken place. The rationale is that amateur users with little knowledge of image acquisition are more likely to use more convenient, highly compressed images ('quality' 50, for example), higher ISO values (to guarantee good exposure, leading to a low 21 dB SNR) and not be aware of implications of sensor selection (8 MPs as opposed to 12 MPs, such as a Nikon Canon EOS 350D).

8.2.2 Geoscientists with knowledge of photogrammetry, but lacking in image acquisition/RAW data handling (Group 2)

This group represents geoscientific researchers, who whilst are proficient at using photogrammetry and generally discuss how data has been acquired, don't provide details on the camera settings and/or image formats the data was captured, stored and processed in. As a result, the research may not be reproducible and so call into question some of the quality of some of the data if care is not taken in its reporting.

For this scenario, images were reduced to 10 MPs from the native 12-MP resolution, a small amount of additive noise was introduced (30 dB additive noise) and the image were JPEG compressed to the default level of 'quality' 92. The rationale is that these users will be experienced in data collection, but be unaware of the potential benefits of capturing RAW images (and therefore using 'default' JPEG quality), have some noise present as a result not optimising the ISO level (30 dB SNR) and not having access to high quality equipment (an example camera would be a Sony A200K, at 10 MP).

8.2.3 Expert users with deep knowledge of image handling (Group 3)

This group represents those who understand the potential degradation associated with using compressed images, and value access to raw data and associated metadata. These users will be aware of how best to capture data in an optimal way and capture in RAW formats to allow for maximum flexibility within image processing stages. For this scenario, TIF images derived from the RAW data files were used, with no downsampling performed on the images and no additional noise introduced.

8.3 Photogrammetric results

8.3.1 Hunstanton

Results from this chapter are presented in the context of the results from Chapter 4, with which they are linked. For the Hunstanton site, the results are largely consistent with what we have seen in the previous chapters (Table 8.1). The error for the first user group is the largest of the three, and the largest of any of the sets from any of the other results presented within this text, aside from the very heavily compressed sets seen in Section 5.1.1. The compressed set from Section 5.1.1, which had no added noise and used the full-sized images, returned a median error of 7.04 mm, lower than the 7.62 mm median error seen for the first scenarios results for Hunstanton. Equally, for images downsampled to 81.5 % of their original size, as seen in Section 5.1.3, error was 7.00 mm, suggesting that these two operations, adding noise and downsampling, degrade the photogrammetric products somewhat independently. Considering this, the results from scenario

1 suggest that error propagates through RAW data from these sources in a manner which combines to amplify the degradation (which is apparent in Figure 8.1 B, when comparing scenario 1 with scenario 3).

In scenario 2, this trend is not seen to continue, as the resulting median error for the combination of minimal noise (30 dB, which had a slight effect reported in Section 5.1.2), downsampling to 91 % of their original size (which had similar median error to scenario 2 of 6.8 mm as opposed to 6.82 mm) and compression was almost identical to that of just the downsampling on its own, albeit with a slightly lower point count than any of those individually. Thus, the conclusion is that, in this scenario, GSD is by far the most important consideration when acquiring data in terms of accuracy.

For the third scenario, the ideal outcome is seen, where point count is highest and error lowest. This would be expected with artificial degradation of the same images, but the scale at which the degradation happens is important to note.

8.3.2 Overstrand

The Overstrand results, as seen within the other results chapters, exhibit no discernible trend in terms of error produced, though point count is seen to fall between scenarios at a comparable level to that at Hunstanton. For the first scenario, where images are degraded to their worst possible quality given the scenarios, error is seen to be the lowest of the three. RMSE-P from the image matching stage is also lower than that of the second scenario, suggesting that the downsampling to 10 MPs has a negative effect at that scale. This was also seen in the downsampling results in Section 5.1.3, where the image set downsampled to 10 MPs also showed the highest median error. At this resolution specifically, both the image matching and dense matching are failing to produce as good results as any other resolution for the Overstrand site. Image quality, as noted previously, does not appear to be a limiting factor for these results.

For the third scenario, RMSE-P is lowest and point count is highest, which is part of the trend seen for the Overstrand site for the results. For the portion of the back wall presented in Figure 8.2, scenario 3 clearly outperforms scenario 1 in terms of both error and density, suggesting the difference is localised to specific area within the products.

Table 8.1 Results of the three scenarios for each site.

Site	Scenario	Mean GCP Error (mm)	RMSE-P	Error (mm)	Point count (millions)
Hunstanton	1	5.5	0.33	7.62	11.37
Hunstanton	2	5.12	0.3	6.82	15.67
Hunstanton	3	5.36	0.28	6.23	19.84
Overstrand	1	10.04	0.46	20.03	12.2
Overstrand	2	9.88	0.64	24.58	17.3
Overstrand	3	8.3	0.39	20.2	22.2

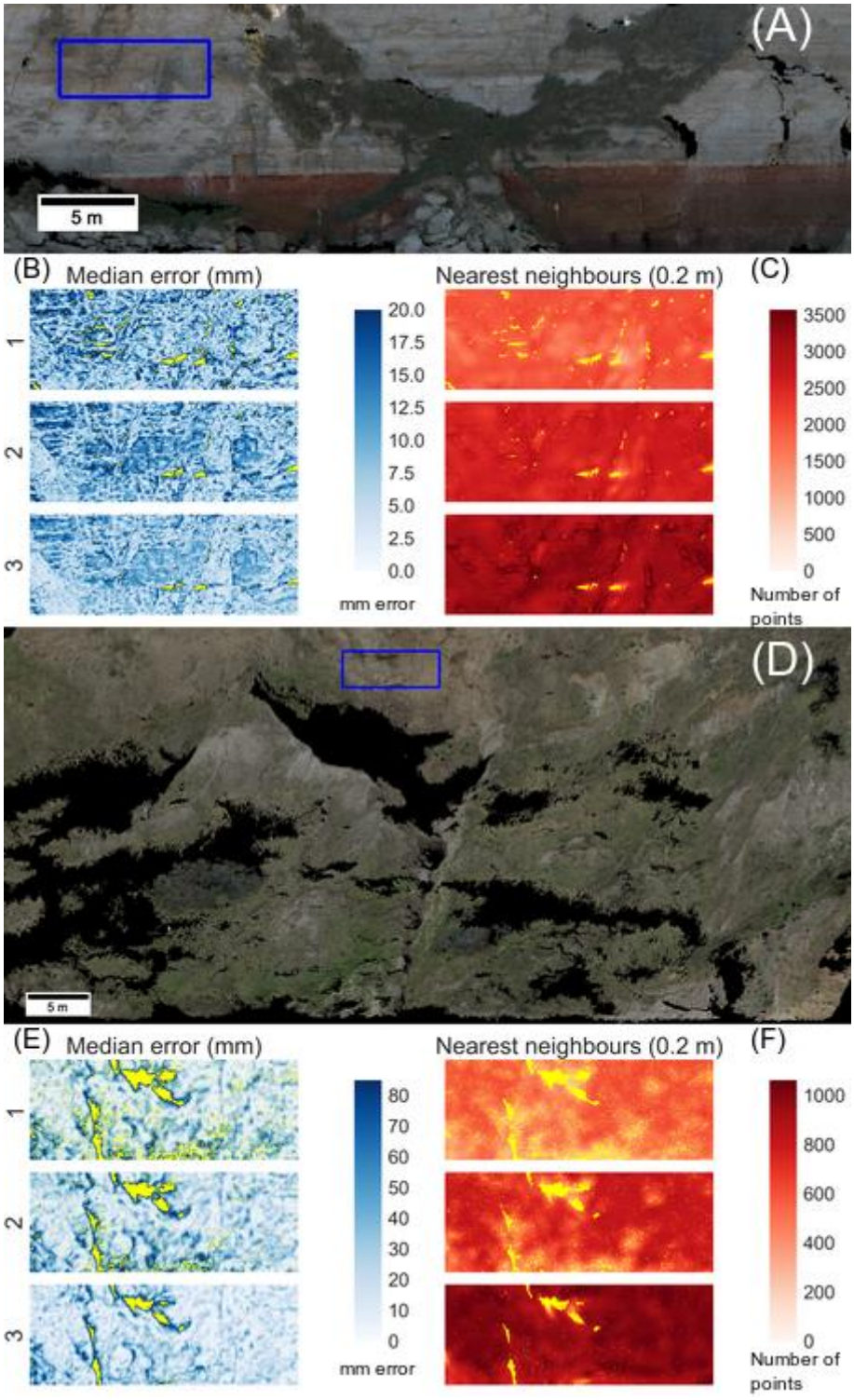

Figure 8.1. Within these three scenarios, mote acute differences are seen at both sites for each product. For the Hunstanton results, error was much higher within the first groups product (B), with much sparser clouds (C) for the upper portion of the rock face. The same trend was observed at Overstrand (E and F) for the back wall of the landslide. No data (missing) values are in yellow.

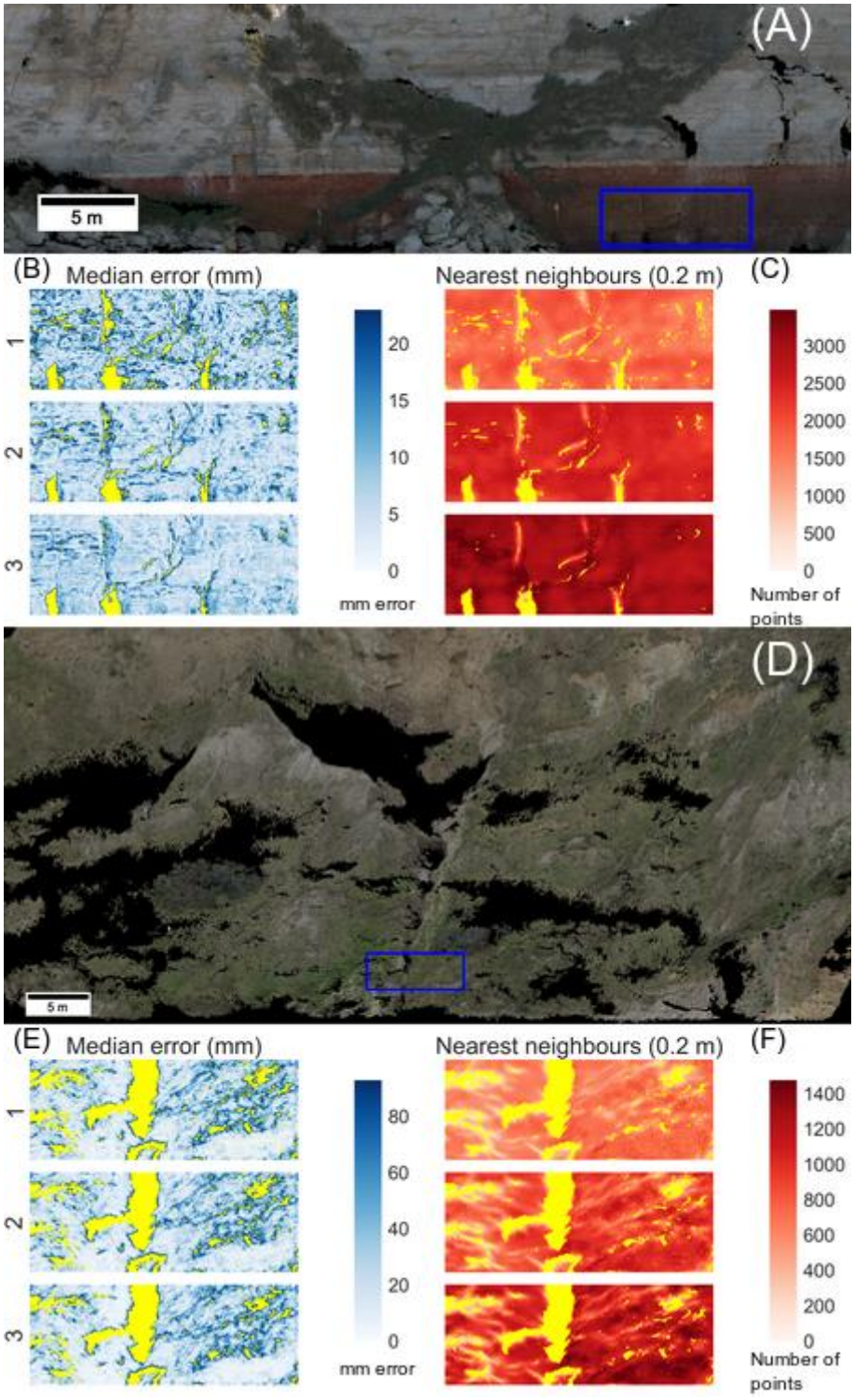

Figure 8.2 Median error is higher with holes beginning to appear in the point cloud for the Hunstanton results under scenario 1 when compared with scenario 3 (B) in the carstone region of Hunstanton. For the Overstrand results, very little difference is seen in terms of error between the three, though some thinning is seen towards the front portion of the landslip for the downsampled scenarios (F). No data (missing) values are in yellow.

8.4 Difference of DSMs

Difference of DSMs for both sites are presented specifically for scenarios 1 and 3, as spatial patterns as a result of degradation are investigated. It is clear from Figure 8.3 that the degradation of the imagery by adding noise, compression and downsampling has had an acute negative effect on the point cloud products generated, with median error 1.39 mm higher. The degradation appears to show no spatial trend, suggesting that image quality is the limiting factor. If a spatial trend were seen, it would suggest that the results may not be valid, as perhaps image matching would have failed in one part of the point cloud.

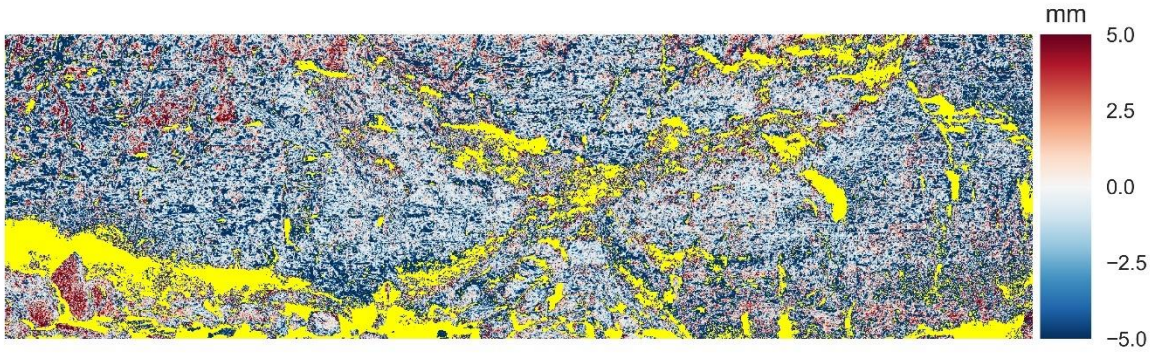

Figure 8.3. Difference of error DSM for scenario 3 compared with scenario 1. Blue pixels represent areas where scenario 3 is performing better than scenario 1. For the Hunstanton results, scenario 1 performs far worse than scenario 3 in every area of the point cloud products. No data (missing) values are in yellow.

For the Overstrand results, a more ambiguous picture is presented (Figure 8.4), with large differences between the results for the front right-hand edge of the rock face, suggesting that this area is a reason why scenario 1 has performed well. The higher resolution images produce more points in vegetated areas due to the increase in resolution, whereas the worse quality images only match points in well textured regions

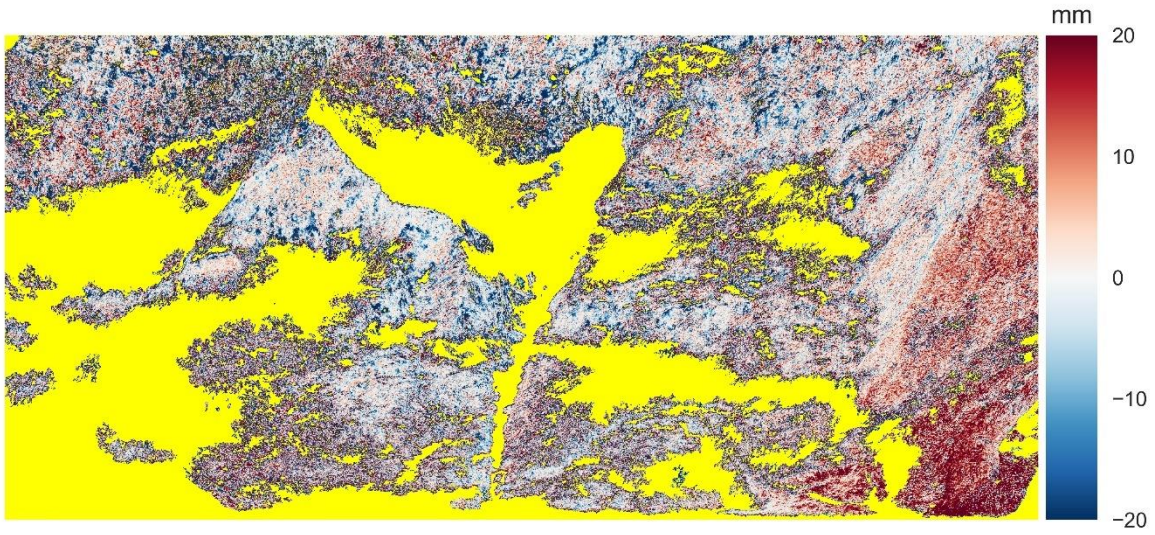

Figure 8.4. Difference of error DSM for scenario 3 compared with scenario 1. Blue pixels represent areas where scenario 3 is performing better than scenario 1. For the Overstrand results, these two scenarios return error at similar levels. No data (missing) values are in yellow.

8.5 Summary

Within this chapter scenarios representing common user groups with varying levels of experience in image acquisition and data handling were presented, to contextualize the results seen within the preceding chapters, as well as the new results presented within this one. The results demonstrate that, depending on the complexity and type of site being investigated, poor data handling could have acute effects on the quality of the product being produced, where simple steps could be taken to mitigate these.

Chapter 9: Discussion

9.1 Introduction

Within this chapter, results generated from each of the previous chapters will be discussed, limitations of the research addressed and recommendations stemming from the results presented. The two sites studied within this work produced markedly different results, and so are treated as separate end members of the workflow presented when it comes to the discussion of their results. The Hunstanton site is a simpler case than the Overstrand site, which represents an extremely difficult site to capture using SfM photogrammetry.

The variation in median error and point counts arising from different image operations are highly variable between the two sites presented, with some apparent trends being seen at Hunstanton which are not reproducible for the Overstrand site. Specifically, while compression, noise and downsampling were all seen to contribute to degradation of the results at Hunstanton, these same trends were not seen for the Overstrand results, except for point count reducing with downsampling. Difficulties within image registration for the Overstrand site indicate that these errors are likely linked with the lack of fixed points the image matching algorithm could detect and match on, meaning that image matching could only occur across a few regions of exposed mud which were static between images. Across the vegetated parts of Overstrand there were few tie points, meaning that the muddy wall towards the back of the landslide were where many tie points were drawn from, making image registration dependent on features in this area. The vegetation is the cause of the poor registration, whereas for Hunstanton enough of the rock cliff was not vegetated making the registration straightforward.

For the Hunstanton site, image registration is better, as the study area was relatively stationary. The image matching and bundle adjustment did not return any misaligned cameras and tie points show good coverage across the site in every image set processed. Image quality is therefore likely a limiting factor at the Hunstanton site, but not at the Overstrand site, where photogrammetric processing could not perform optimally due to moving objects (such as vegetation blowing in the wind) within the scene. Thus, many of the recommendations to do with image quality for users are drawn from the results of the Hunstanton site, where results from the Overstrand site exemplify how users should be addressing scenes with complex scene geometry.

9.2 Limitations of this research

The results presented have some limitations, including those to do with the accuracy of the ground control points, accuracy of reference surface, RAW image quality and photogrammetric processing.

Firstly, the weather was notably different between sites. For the duration of the day at Hunstanton, the sun was mainly overhead, though heavy winds blew down one of the photogrammetric targets during the total station survey, meaning it could not be used within the experiments. Towards the end of the photogrammetric survey, the sun was behind the cliff to the East. The bright sun caused exposure metering to underexpose the area which was in shadow. Ideally, light would be diffusely spread across a scene (i.e. cloudy), eliminating shadows. In addition, lens flare shows blue patches on the image because of internal reflection of the lens elements, which is represented in the RGB point cloud by a patch of blue points on the cliff face (Figure 9.1).

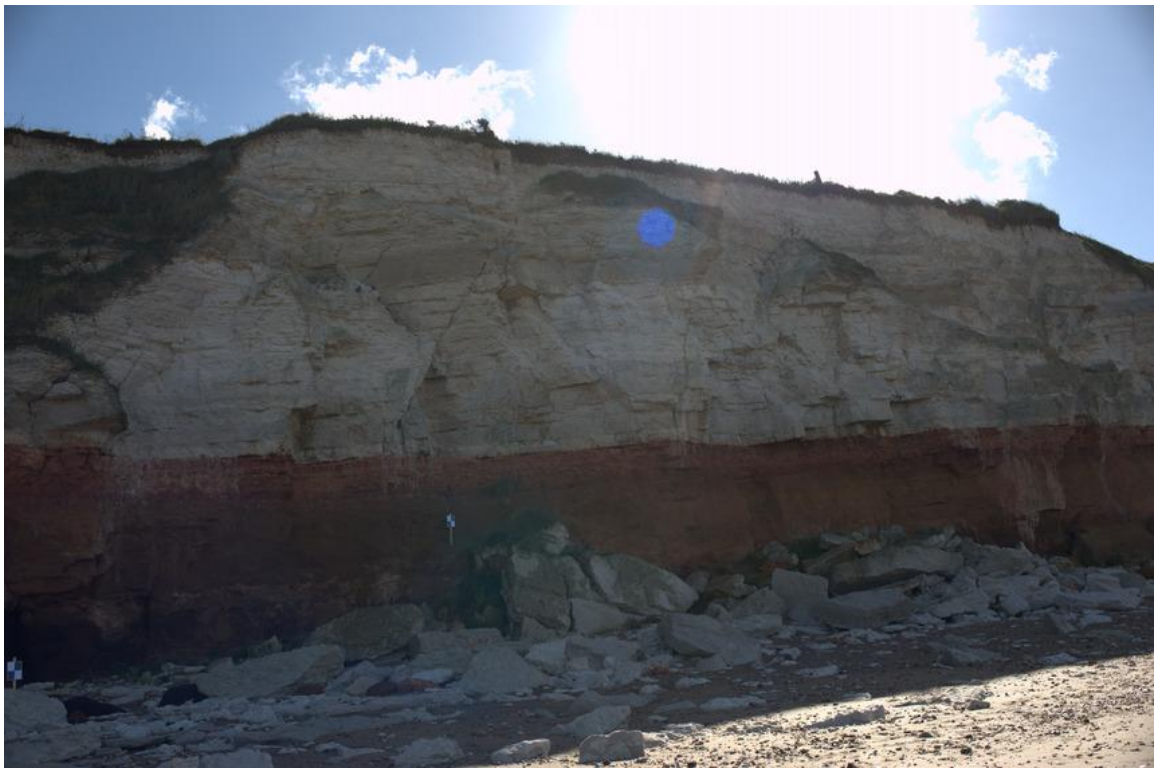

Figure 9.1. The sun was setting behind the cliff towards the end of the photogrammetric survey at Hunstanton, meaning some of the images contain lens flare and less than optimal exposure. Image number 35, facing south-east, taken at 13:31, ISO 200, aperture f/8, shutter speed 1/250 s.

The wind was blowing strongly enough to cause movement of the camera on the tripod, which was a source of blur within the imagery. This is not perceivable when viewing image thumbnails,

or even when zooming in to crop sections. Nonetheless, a small amount of blur likely exists within each image, which could have been a cause of image degradation across the survey. This is difficult to quantify objectively, as blur detection algorithms are not ubiquitous. Some attempts have been made to perform this automatically (Sieberth et al., 2013), though this was not applied in this study as these algorithms are still relatively nascent. The LiDAR scanner used was heavy enough to not be influenced by the wind, and the targets were straightforward to locate and map within the total station survey.

During the photogrammetric portion of the fieldwork at Overstrand, rain and wind interrupted surveying. This degraded imagery that was acquired, as water dropped on to the camera lens. As a result, the imagery was acquired later in the day during different weather conditions. Conversely, the total station survey and LiDAR scanning were undertaken at the beginning of the day in stable weather conditions. As noted in 3.3.2, the LiDAR scanner could not locate one of the targets accurately, and so only three targets were used to coregister the LiDAR survey. This brings its accuracy as a reference surface into question, as the registration of each laser scan onto one another has residuals of 2-3 mm with no redundant points for verifying this accuracy. The manufacturer further notes that the expected 3 mm one sigma error per point is assuming a 78 % albedo. For the mud banks at Overstrand, this was likely not the case, as the mud was dark, introducing a further unquantifiable source of error.

While results presented focus on the median error, 75-80 % of the points returned fall within the 3 mm one-sigma error reported by the TLS manufacturer at the surveying range for the Hunstanton site. Thus, while increases in these median errors suggest an overall degradation in the quality of the product, the accuracy of the TLS acting as a reference should be kept in mind when discussing the results. In addition, the residuals between the TS survey and the TLS registration (5 mm for Hunstanton and 10 mm at Overstrand) represent a similar scale of error when compared to the median errors included within the results. Thus, while the objective of the research was to confidently state that image degradation has a demonstrable effect on the accuracy of photogrammetric reconstructions, the conclusion is that image degradation has a demonstrable effect on the accuracy of reconstructions in the context of comparison with the reference surface generated.

However, the increase in median error with image compression, additive noise and downsampling is consistent at Hunstanton, which gives confidence that the trend is valid. Limitations, are summarised in Table 9.1, with improvements to the research design on how to mitigate these included, should the research be repeated.

Table 9.1 Limitations of the research, with recommendations on how to improve upon the research design

Limitation	Recommendation
Accuracy of TLS is on the same order as GSD	Use a denser reference scan, with a higher accuracy laser scanner. Survey closer to the surface in question.
TS surveys have a relatively large RMSE	Acquire coordinates of targets from more than 2 stations to allow for greater redundancy
Overstrand camera positions had to be initialized	Review survey camera positions and pose, consider using more reference targets to allow easier image registration

9.3 Photogrammetry is tolerant to JPEG compression

For the Hunstanton survey, results suggest that increasing JPEG compression does degrade the quality of the point clouds and DSMs produced from the photogrammetric software. However, many of the compressed image sets return acceptable results in terms of median error when compared with the uncompressed image sets. For compression ratios of 20:1, representative of JPEG compression used in Nikon's 'basic' image quality setting, the median error of the point clouds increase just 14% when compared with the uncompressed images, and just 8.5% over the image sets with JPEG compression of 'quality' 92, the typical level of compression associated with off-camera JPEGs, which represents a compression ratio of 4:1. Compared with the amount of variation in error which may be expected within photogrammetric processing due to stochastic processes from Section 4.2, this 8.5% represents over twice the range of error seen within identical image sets over 10 runs of the workflow for the Constitution Hill pilot study, and over 20 times the range of error within identical image sets over 10 runs of the Hunstanton data. This indicates that the results from the high quality JPEGs are showing demonstrable degradation when compared with the TIF images.

A good example of a study which could have used this granularity of reporting is Vasuki et al. (2014), who used images to execute SfM (using VisualSFM) and fault line detection algorithms of a layered meta-sedimentary sequence cross-cut by a series of dikes and faults. Given the similarity in texture between their study site and ours, image quality is likely to be a limiting factor in the

accuracy of the SfM products, but the authors do not report on file formats in the related paper that details the image acquisition strategy (Micklethwaite et al. 2012). This probably had an impact on the quality of the results for both the SfM and fault line detection workflows which went unreported, and likely led to sub-optimal results.

For applications where point cloud accuracy is not required at the highest levels, such as highly vegetated areas where high quality reproduction is impractical owing to the lack of fixed points, more compressed data could be used to generate results accurate enough for that application. Additionally, for longitudinal studies where storage capacity might become an issue, for example time series generation at remote locations using automatically triggered cameras, compression can be applied at the point of capture if required.

For the Overstrand results, the trend of additional compression increasing error is not held. The image set compressed to 'Quality' 50 performs better than the uncompressed images for this metric. This suggests that at Overstrand, image quality was not a limiting factor, and the accuracy and completeness of the resulting point clouds were more likely a result of the survey design, complications with image registration or accuracy of the point cloud reference. Considering the intrinsic error within the TLS and the difficulty of the scene for photogrammetric processing, this is not surprising, though the fact that little to no effect is seen with changing image quality suggests that for casual and/or scoping surveys in difficult areas, image quality should not be a limiting factor for collecting initial data. For ecological applications, where photogrammetry is being used to generate volumetric estimates of bush extent (Cunliffe et al., 2016), results suggest this is a reasonable methodology which is not image quality limited, so low-cost cameras on light UAVs could be used to give rapid overviews of study sites.

9.4 Image downsampling thins point clouds proportionately to megapixel count and increases error

Downsampled images were used to investigate how spatial sampling frequency interacts with both accuracy and density of the point clouds produced. For the Hunstanton results, downsampling affected both the accuracy and density of resulting point clouds, with density dropping dramatically as the images were downsampled. It is noted that densification is a metric which is not directly linked with image quality, as upsampling of data can increase densification, and so should be considered with respect to the number of pixels per image. Differences in point count are in line with the ratio of the image size, and so density is seen to be reasonably uniform through the sample. This suggests that pixel count is the biggest driver of point count, as results

show that adding noise and compressing the images (except for high levels of compression) show little effect on the point counts at Hunstanton.

Interestingly, pixel count after downsampling does seem to affect the median error of the point clouds produced for the Hunstanton results (Figure 9.2). This is intuitive, as pixels, and thus points, represent a larger area on the ground. For example, this could be due to the size of the features being recovered: if there is micro topography on a boulder due to cuts into the face, generalizing the pixels will lose the small amount of topographic information, which will lead to a slight increase in error due to recovered points being on a plane with the rock face instead of being set back. Also, due to sampling sparsity, there is a larger chance a point will be located farther from the reference set. Considering GSD in the original size images is 7 mm at Hunstanton, the increase to 14 mm effective GSD is likely contributing to the degradation.

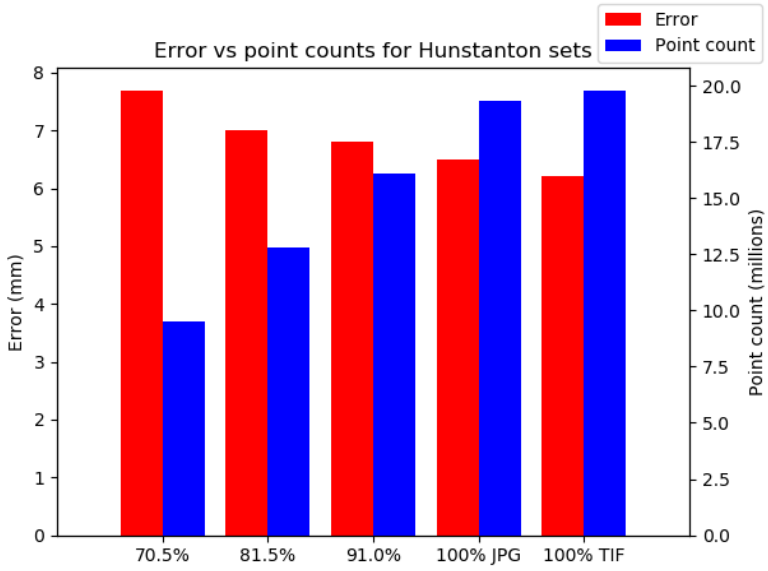

Figure 9.2 The Hunstanton point counts increase dramatically as size of the images increases, though the median error is also seen to drop, albeit less dramatically.

RMSE-P shows no notable differences between images downsampled by different amounts. This is surprising, as it would be expected that the downsampled sets would show a lower RMSE-P, as bigger, more general pixels would aid the image matching process in terms of pixel error. For example, where one image contains pixels which are noisy, a suppression of this noise might be expected when downsampling is applied, though this is not the case. The results at Hunstanton show low RMS error, so any slight difference could indicate a proportionally large difference in the quality of image registration.

For the Overstrand results, a similar trend in point count is observed. The number of points within a photogrammetric reconstruction is proportional to the size of the input image, though within the image set downsampled to 81.5 %, a greater than proportional drop in density is seen. In terms of error, there is no relationship between image size and RMSE-P, though two sizes of images (100 % and 81.5 %) return lower median errors than the other two image sizes. This suggests that generalizing pixels through downsampling might be effective for optimising surveys where image quality is not a limiting factor.

The results between sites again suggest that the Hunstanton images are limiting the quality of the product being returned whereas for the Overstrand results, due to the difficult survey conditions, image quality doesn't influence the accuracy of the results. Where image quality is a limiting factor, sensor size should be maximised to ensure the best quality results. Where it is not, smaller sensors can likely be applied to return comparable results, if care is taken with the survey design. Where issues with complex geometries could be addressed by a higher sampling rate, pixel pitch and sensor size should be given a priority in survey planning.

An example of a study where these results are relevant, is that of Westoby et al. (2012), whose data was used within one of the pilot studies. Within their study, they report that images were 're-scaled to 55 % to reduce computational demand'. Whilst they justified this computationally, the point was never revisited, and likely had an important effect on the results being reported. This makes comparison with other studies more difficult, as while the sensor used is reported (Panasonic DMC-G10 12 MP), no attempt to quantify the impact of the reduction of image size was made, making like-for-like comparison more difficult. Our results, however, provide a framework for investigating how this downsampling would impact results, to make this comparison more streamlined.

9.5 Adding noise degrades blocks, SNR should be reported with metadata

Adding variable levels of noise to the images was intended to demonstrate any effect that this had on the image products. Whilst there are several noise sources present when acquiring an image, as discussed in Chapter 2, for the purposes of these tests Gaussian noise was used.

For the Hunstanton results, adding noise had a subtle effect on all of the RMSE-P residual from image matching (ranging from 0.28 – 0.66 pixels), median error (ranging from 6.22 mm to 7.37 mm) and point count (ranging from 16.4 million points to 19.8 million points). This suggests that where image quality is a limiting factor, sensor performance is required to be reported with the image metadata to establish a baseline. For comparing surveys using different cameras or surveys with a common camera but different settings, including an estimate of SNR is required to ensure

sensor performance is accounted for. This should be considered independently from lens performance or degradation due to camera motion or occlusions. Benchmark testing can be accessed using a variety of different databases, including but not limited to DxOMark, (2018) and LensScore, (2018).

For the Overstrand survey, RMSE-P increases as with the Hunstanton results, though this increase in image matching error is not reflected within the median error of each product. This indicates that at a pixel-by-pixel level, features being detected are more difficult to match, though when running the dense cloud algorithm the solution does not appear to be adversely affected by these inaccuracies within image matching. Point density falls with increasing noise, indicating that some degradation is being carried through to the dense cloud as with Hunstanton, though this is only slight for the Overstrand site.

The results indicate that users should aim to keep ISO to as low a value as is practically possible. While higher ISO settings are useful if the camera is not tripod mounted, or environmental conditions introduce the risk of blur, users should be aware of the trade-off of this decision.

9.6 Single channel conversions only produce slight differences

For the greyscale results at Hunstanton, little difference was seen between any of the point cloud products generated using different bands or greyscaling techniques. Differences in each RGB combination showed that physical effects, such as the focal planes of each colour band being different, were compensated for. Within the non-linear techniques, little difference and no improvement on the point cloud products for median error and point density were seen. Whilst PCALab produced the lowest RMSE-P of any of the results for both sites, they were only marginally better than the RGB images.

For the results of the lens models from the Hunstanton survey, some asymmetry was seen. When fitting just the blue band, the focal length on the X-axis (35.942 mm) is slightly longer than that of the y axis (35.926 mm). This represents a 16 µm difference between x and y axes, or 2 pixels. This asymmetry in focal length balances the green and red bands, apparently more evident in the blue channel.

The fact that the lens models are compensating for slight differences in focal lengths due to the three bands falling on slightly different focal planes shows the sensitivity of photogrammetric methods. This trend is seen to be broadly true, with the individual blue band an outlier in the sense that it's focal length is shorter than that of combinations even where it features as a large proportion of the greyscale band, and the fact that the fitted focal length shows discrepancy in x

and y. This suggests that the blue channel may be performing worse than the other channels due to differences in how light is focused by the lens, though this could also be an outlier because of the stochastic nature of bundle adjustment.

For the Overstrand lens model results, this blue band outlier was not evident, being close to the green channel in focal length space (as described in Section 3.4.5), and there being no evidence of asymmetry in the pixels/focal lengths across any of the combinations presented. This is likely due to the requirement of the image sets to be initialized with camera positions due to consistent errors in the image matching stages of the workflow, as well as the differences in colour between sites. Image positions were static during the entirety of the Overstrand experiments, and thus small optimisations in camera positions were not permitted during bundle adjustment, as opposed to Hunstanton where this was the case. The fact that the dataset shows a less consistent trend, in terms of those greyscale channels made up of a combination of bands, is likely a result of this difficulty.

Considering the same camera was used for the tests at both sites, these two datasets are comparable. A similar increase in the focal length, as set during bundle adjustment, for combinations of bands representing longer foal lengths were seen for both sites, though some noteworthy differences include the fact that at Overstrand, the parameter for focal length, particularly Fy, is larger than at Hunstanton. This may have impacted the image sets' performance relative to the others, as when we move towards longer wavelengths we see a consistent solution found. This indicates that the actual principal point of the lens is closer to the red channel than the blue, as this higher variability suggests the software fits a consistent solution when using red-weighted greyscale images.

For the Hunstanton results, the green band returned the lowest RMSE-P of each of the primary colour channels (Red, Green, Blue). This is intuitively the case, as the green channel is sampled more frequently (twice as often on a Bayer array) than either of the other two. However, we only see this reflected in the point density of the recovered clouds, as error is increased in comparison with the blue band. This suggests that additional points being added, when compared with the blue channel, are of lesser quality. Comparing this with the Overstrand results, a consistency is seen between the sites, with the blue channel returning the lowest error but also less points than the green channel. Selecting the blue channel therefore appears to have an advantage in that spurious points appear to be excluded more frequently.

The spatial patterns indicate a potential interaction between the distance the scene features are from the camera, and the error associated with that area. Scene features towards the back of

each site are better reproduced in the red-weighted channels whereas for blue weighted channels the opposite was seen to be true. This is likely a result of the lens model fitted.

With regards to the further greyscale transformations presented, PCALab, produces the most consistent results, with the median error close to the RGB input at Hunstanton and lower than that at Overstrand. The recommendation from these greyscaling experiments is that because only slight variation within the error is produced, these techniques are likely confined to scenes where colour varies more widely than that seen at the study sites within this work. Where expert users might find value using greyscaling techniques for specific applications or sites, in general they do not have a large effect on the accuracy or density of these photogrammetric products.

One alternative to using image processing to attempt to maximize accuracy of the data from colour information is to not use a colour filter array. Monochrome cameras, such as the Leica-m monochrom (Leica, 2016), have no colour filters overlain on the sensor. These could be used to maximize spatial resolution and avoid the need for interpolation of each colour channel. Another benefit is a higher per-pixel sensitivity because of no photons being filtered before arriving at each pixel.

The results of our studies show little variation with the single channel generated, with the best single channel set performing just 0.5 % better than the untreated RGB images in terms of median error at the Hunstanton site. A limitation of this study, however, is that only two sites were investigated, so these conclusions would benefit from further analyses of this nature.

To extend this research, processing subsets of the images with different greyscale treatments based on the content of each image could offer the opportunity to fine tune each greyscale image based on their contents. Another would be (eg. Westoby et al. 2012), to execute the SfM workflow on these subsets independently, and recombine each point cloud at the end of the workflow. In the context of other research, however, our findings show little return for the overhead of pre-processing imagery to optimise the single-channel.

9.7 Photogrammetric blocks containing more images produce denser clouds

The results show that photogrammetric blocks with fewer images produce fewer points, though this trend shows some interesting exceptions. For the Hunstanton results, set 7 (Table 6.1) contains widely spaced viewing stations (number 1, 4 and 7), with a lower point count when compared with image sets where fewer images were captured from more positions. For example, set 6 contains just 8 images but returned a higher point count than set 7, suggesting that more frequently spaced images at fewer viewing angles perform better than wider spaced images with

more viewing angles. In terms of error, a lower median error in the block with fewer images is seen. In set 6 a low RMSE-P (0.09) suggests it performs well in terms of both image registration and dense cloud generation.

These results were similar for the Overstrand site, though set 6 shows a higher RMSE-P (0.266) when compared to the wider spaced set 7 (0.215). However, as seen with the Hunstanton results, both the median error and point count perform better for set 6 . These indicate that even with a different GSD, scale and scene features, error and point count show a tendency to perform better where fewer images were taken closer together. The consistency between sites indicates confidence in these trends, though this is not universal. The lowest median error (13.29 mm) for the Overstrand results was seen in the image set containing the fewest images, indicating that adding images did not improve performance in terms of this metric, though relatively few points were returned (4.19 million, as compared with the full set, which returned 20.44 million points). Notwithstanding, it appears that taking images at slight angles (15 °) from the surface normal with 80 % nadir overlap is a good strategy for generating useable results with limited images. This is corroborated by literature suggesting incidence angles be kept to <20° from the surface normal (Mosbrucker et al., 2017). In this work, sets 2, 5 and 8 broke this rule of thumb, and produced poor point density compared with the other sets.

The results confirm that carefully planning surveys in the first instance is valuable, as this will have effects on the quality of products independent of the factors discussed within the other chapters. When the number of images which can be acquired in a given survey is limited, a balance between reasonable amounts of image overlap (such as 60 % overlap at Nadir for set 3) and convergent network geometry (at convergent angles <20 °) to mitigate systematic error appears to work best. This is particularly acute in aerial surveying, where platform flight time can be limited, and so the flight path needs to be configured so that data are acquired from optimal locations. For instance, a circular flight path with angled cameras whose optical axes converge would be preferable to a transect flight path with a nadir pointing camera. This is in line with findings in the literature, which show a large degradation in the accuracy of photogrammetric products when convergence is > 30° (Stumpf et al. 2015, Mosbrucker et al. 2017).

James et al. (2014) demonstrated this further using simulations of typical UAV acquisition geometries. They show that ‘weak’ imaging geometries, where sufficient overlap isn’t present and convergent imagery isn’t optimized, cause significant degradations to the precision of photogrammetric products. Conversely, Wackrow et al. (2011) showed that convergent angles of 10 ° can aid in self-calibration of the lens model and reduce dishing/doming. These insights are

reflected within the results of Chapter 7, where set 3, which has images overlapping at 15 °, performs well for the Hunstanton site (median error of 6.65 mm with 17.6 million points). Set 2, containing 50 % more images (24 as opposed to 16) returns lower median error to set 3 (6.41 mm), though less points (17.1 million), even with the additional images.

The benefit from densifying the camera network, in absolute terms for both error and point count shows diminishing returns, with the most complete photogrammetric block (set 1) returning similar results to both sets 2 and 3, which have 60 % and 40 % of the images that the full block has. With just 40% of the images, set 3 returns 90 % the number of points recovered by set 1 at Hunstanton, which could be useful information for studies where only a limited number of images can be recovered for practical reasons (eg. longitudinal studies where the cameras will be stationary), as well as aerial surveying where flight duration is limited, in poor environmental conditions where only a limited number of locations can be visited.

Another notable difference between the two sites is the rate at which the point count decreases as images were removed from the original set (set 1). For the Hunstanton results, the cloud with the least number of points returned 56% the number of points when compared with the full block, while at Overstrand the thinnest cloud returned just 20.5%. This disparity could be caused by a difference in scene features. As discussed previously, the Overstrand site represents the upper limit of what is practically possible to capture using photogrammetry in terms of complexity, given the difficult scene geometry and content. The extra images are invaluable in localising the finer detail within the scene (vegetation on the banks for example). When removed, the algorithm suffers more at Overstrand, with its fine detail and moving scene features, than the relatively stable Hunstanton. It therefore appears to be better to take more images in conditions where the scene features are challenging to ensure a reasonably dense cloud is recovered.

In terms of processing time for each reconstruction, there were large differences between the sets presented. Acute examples showing the effect of introducing redundancy are at Hunstanton, (using the settings given in Section 3.4.5); set 3, containing 16 images took 45 minutes as opposed to the full 40 set of images taking 270 minutes on a single computer with an Intel 4770k CPU, 16 GB of RAM and an Nvidia GTX770 graphics card. In situations where timely information is required, this could influence the density of the network required for a given task. A second example from Overstrand was set 9 (6 images) was completed in 4 minutes when compared with the full image set (40 images) which took 240 minutes. For scenarios with complex geometry, where densification does not have a significant impact on the accuracy required, this should be considered depending on the use case. For example, where a time series is required, the images

could be acquired, data processed and results stored in a single day, whereas it would be impractical to do this with a full image set.

9.8 Stochastic effects make results even more uncertain

While a range of topics have been discussed within this chapter that could explain the results observed, conclusions are more complex due to the stochastic nature of SfM processing. Ideally, each of the model runs should be repeated several times to research whether differences in error between the products are of a similar order to differences between one model run multiple times. In lieu of this, the pilot studies' results indicate that for a similar coastal study, given similar scene features, similar control and similar image quality, the results are stable when run multiple times (Section 4.2). This is corroborated by multiple runs for the Hunstanton dataset, which showed low amounts of variation (0.14% of the mean) over 10 runs.

The smoothness of the transition within the error triangle for the Hunstanton results (Figure 5.3) suggest a low standard deviation between clouds made up of similar channels, and smooth gradients across the triangle. This is not the case at Overstrand, where the error triangle is relatively heterogeneous, with peaks existing next to troughs. This indicates that Overstrand is likely more sensitive to initialization conditions than Hunstanton, and as such conclusions drawn from its products are likely less reliable.

The low variance in similar single channel TIF sets at Hunstanton does indicate that the results are stable, every one of which performed better than the 'quality' 92 compressed JPEG sets from Chapter 4. Thus, the conclusion is that the benefit of using images with a greater bit depth is evident compared to JPEG compressed images, as is the conclusion that the results presented are stable enough to draw comparisons between sets independent of stochastic variance.

Recent developments within the field have attempted to account for inaccuracies which arises from both the Chamfer matching distance and lack of precision estimates within the bundle adjustment stage. Lague et al., (2013) introduced the multiscale model to model cloud comparison (M3C2) tool which better accounts for internal variability within each point cloud when comparing two clouds by calculating a confidence interval for subsamples of points based on local roughness surrounding the given point. They note the weaknesses associated with the Chamfer matching distance, namely the sensitivity to the cloud's roughness, outliers and point spacing and how M3C2 deals with this by accounting for roughness and point spacing while omitting outliers. Barnhart et al. (2013) compared the two techniques in their study of a permafrost degradation feature and concluded that M3C2 was more robust as it accounted for surface roughness, as well as registration uncertainties.

M3C2 likely could have better handled some of the difficulties in cloud comparison at the Overstrand site, which had large gaps and roughness variations due to the scene contents. However, at the time of the research within this thesis, the M3C2 algorithm was not accessible in CloudCompare by command line, and so could not be built in to a standard, scripted workflow.

More recently, James et al. (2017b) introduced the concept of 'precision maps' to the SfM community, which captures precision information regarding individual points involved in the sparse point cloud generation, as well as parameters within the camera model. This is achieved by repeating the bundle adjustment step of the SfM workflow many times and measuring the variability of each point involved in the bundle adjustment. These precision maps, which are photogrammetrically-relevant, can be used in place of the local subsample roughness measurements within the M3C2 algorithm to improve initial estimates. In the context of this thesis, these precision maps could have helped contextualize the error metrics calculated by presenting a precision map alongside.

For example, precision maps could be generated for both high and low quality imagery, and these maps compared to visualize where precision is being lost in measurements of individual points due to image quality, as well as the spatial distribution of this precision loss. This is a logical and exciting next step in this image quality research.

9.9 Errors propagate to degrade results

The results from Chapter 7 suggest that the degradation in quality of photogrammetric products resulting from each image quality operation (noise, downsampling and quality) combine to degrade the results more than any operation independently. This is evident in the results from set 3, which was downsampled, noise added and heavily compressed, producing poorer results than any of these operations individually. For images acquired from the same positions, reasonable levels of degradation which would be expected using smartphone or low-cost compact cameras, a 22 % difference in median error was seen for the Hunstanton results. For people new to the field, or those practitioners using SfM regularly as a primary data source, these outcomes could be presented as the same in terms of methodology and image acquisition, but the results show that large differences in image quality can go unreported.

This indicates that there is value in thoroughly reporting image acquisition strategies within methodology sections of published research, and standardising how sensors, lenses and image formats are reported in photogrammetric studies. This is a key step to ensuring that SfM experiments are directly comparable, so that correct decisions can be made on how to capture, store and process data when designing an experiment. Depending on the requirements of an

application, this loss of accuracy could be deemed acceptable, but for the community to accept it, it must be reported correctly.

9.10 Main outcomes and recommendations

Given the results described, several recommendations can be made as to how practitioners of SfM should approach experimental design, and how images should be reported in the literature. Here, these recommendations are summarised, with an optimal form of reporting suggested going forward. Figure 9.3 presents a Venn diagram summarising all image quality considerations to engage users in how to approach these studies.

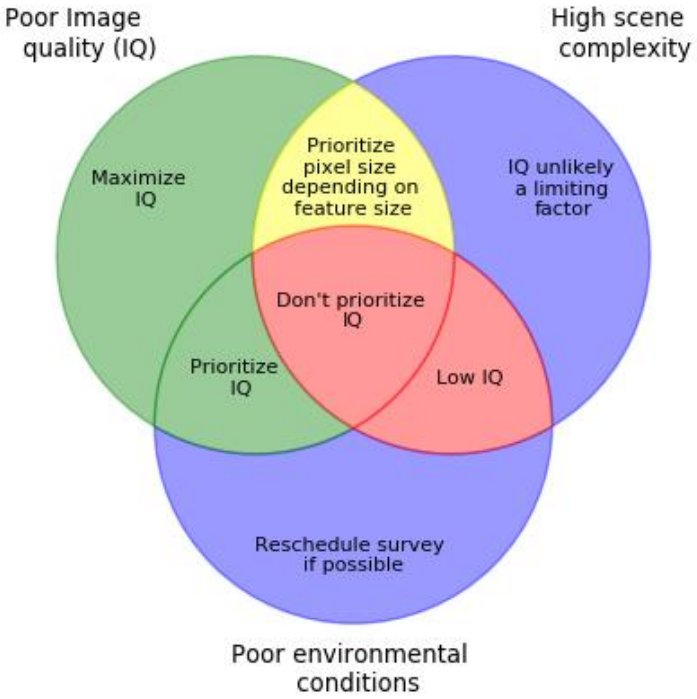

Figure 9.3. Where image quality is a limiting factor, sensor size should be at the largest practical size. Where the photogrammetric product is limited by scene complexity and/or environmental conditions, increasing image quality may offer diminishing returns, as it will likely have little effect on the accuracy of products. If pixel size can be practically reduced to allow a high enough sampling resolution depending on the surface topography, this should be prioritized to capture changes at or near the sampling resolution, if required by the survey.

9.11 Future work

This contribution has aimed to lay the groundwork for how image quality and accuracy should be approached and reported within the geoscientific community. In the future, several extensions could aid in ensuring this is enforced with the rigour and consistency which will allow SfM studies to be directly comparable and reproducible.

Making RAW image data freely available would allow for acquisition practices in experiments to be scrutinised, and potentially for image quality algorithms to assign an IQ score to image sets used in photogrammetry. Having data repositories with benchmark experiments would also allow researchers to familiarise themselves with best practice and so avoid poorer quality data being reported within the literature.

Accurate reporting of image metadata within the scientific literature would also aid in the comparison of studies. Currently many studies have not included sufficient metadata to allow experimental reproduction (appendix O'Connor et al., (2017)). This would be an easy step for researchers to take to allow the intercomparison of results.

The synthetic results presented in this thesis represent a starting point for how one would attempt to compare image quality between studies. Ideally, these data would not be synthesized, and image metadata from a variety of studies be integrated to see if any tangible relationship exists, as described within this contribution. An ideal result would be a prediction of how much of an influence image quality would have on a particular survey, in order to help researchers make better decisions in terms of the equipment they use, acquisition strategies they take and products they deliver. This will undoubtedly be aided by recent developments such as M3C2 (Lague et al. 2013) and precision maps (James et al. 2017b).

Most of the work in this thesis is around the effects of how degrading image quality impacts upon derived photogrammetric products. However, the impact of enhancing photographic products through various techniques, such as exposure stacking (Gómez-Gutiérrez et al. 2015) and focus-stacking (Kontogianni et. al 2017) which aim to enhance individual image quality through blending multiple images into one higher quality image, remain understudied within the geosciences.

Lastly, this work does not extend to the full depth of the data which was captured during the fieldwork. As noted in Section 3.4.3, repeat images and different exposures were captured at every position from which imagery was acquired. Expanding on this, noise reduction algorithms based on median pixel stacking could be used to ensure noise is kept to a minimum. Additionally, for scenes with large ranges of brightness, exposure stacking could be used to enhance images

during pre-processing. These avenues currently remain underexplored but could add value to an expert user's surveys.

Chapter 10: Conclusion

Within this work, image quality, pre-processing and survey design were studied to investigate the effects each of these had on photogrammetric survey accuracy for to sites, a coastal cliff at Hunstanton, UK and a vegetated landslide at Overstrand, UK. For the Hunstanton site, obvious trends are presented which show image quality correlating with accuracy of the photogrammetric survey. The effects of additive noise, image compression and image downsampling operations had reduced the accuracy of the derived products. Pre-processing greyscale images showed no impact on the results produced, and the camera networks with convergent viewing geometries produced the best point clouds in terms of accuracy.

For the Overstrand site, which represents a more difficult site to capture photogrammetrically, a more ambiguous picture is seen. The correlations between adding additive noise and compressing images degrading derived products were not apparent. Pre-processing the imagery through greyscaling did not have any tangible benefits. Densifying the camera network increased the relative number of points when compared with the Hunstanton results, though an increase in accuracy was not seen.

In summary, Hunstanton gives confidence that where image quality is a limiting factor, reducing image quality will acutely impact photogrammetric products. In difficult conditions or at difficult sites, image quality may well not be the limiting factor, so these trends may not be seen.

10.1 Insights from the research

10.1.1 Images can be JPEG degraded where maximum accuracy is not required

For most applications undertaken using SfM, the marginal gains from using RAW images are limited and should be offset against the overhead of the work required. For example, the JPEG images for the Hunstanton study produced a median error of 6.49 mm as opposed to the 6.22 mm for from the RAW images. It is useful to note that in situations where fine detail is required, such as precise estimates of volume, an improvement was seen from using the RAW images. Thus, RAW images should be collected where possible. While there are many situations where they won't be needed/collected, these should be justified by the researcher to establish that this difference in methodology exists.

The recommendation from this research is to only use image compression in scenarios where it is unavoidable. Where RAW imagery can be acquired, this should always be opted for.

10.1.2 Downsampling degrades results

Whilst point density has been shown to be linearly related to image size (for both sites the point counts falls when using smaller images), downsampling has an acute effect on the accuracy of the produced clouds in terms of error (the full size JPEG images produced a median error of 6.49 mm, as opposed to 7.7 mm for the images with half the number of pixels). Users should therefore maximize the number of pixels on a camera sensor for a survey, bearing in mind the impact this could have due to the softening effects of diffraction resulting from the smaller pixels.

The recommendation from this is to use as large a sensor as possible when acquiring imagery and to not downsample imagery prior to executing the photogrammetric workflow.

10.1.3 Noisy images degrade results

From the results and discussion of the image quality chapter, pre-processing and simulating noise degradation of the image data had an impact on the results generated. Adding noise increased error for the Hunstanton results (6.22 mm median error for the images with no additive noise as opposed to 7.37 mm for the noisiest images). SNR performance of the sensor used is rarely reported in published applications of photogrammetry in the geosciences.

Researchers are encouraged to report an estimate of the SNR of the sensor used, along with other relevant image metadata in order to better understand image quality, and the community can move towards using standardised metrics. These results show a spatial pattern related to high contrast changes, so extra care should be taken to acquire the highest quality images where shadows are present, the surface is dark or where little texture exists.

The recommendation from this is to include full metadata of imaging settings in the methodology. This will allow the image noise to be estimated, as well as increasing the reproducibility of the work.

10.1.4 Pre-processing imagery has a limited impact on results

Within the greyscaling chapter, no strong trends emerged in how pre-processing the data could improve results generated from the photogrammetric workflow, with median error for the Hunstanton results ranging from 6.185 to 6.5 mm. As a result, further work is required to determine if differences in results were due to physical effects, such as lens models compensating for chromatic aberration or sampling frequency. Whilst the blue image band did appear to perform better than any of the other colour channels in terms of error, it could not be established whether this was because of specific spatial trends, or points being excluded due to conditions in

the blue channel. At both sites, the green channel producing more points than either of the other primary channels suggests the higher sampling rate is reflected in the products generated.

The recommendation from this is to only experiment with colour pre-processing where colour contrast is high, and there are higher potential gains as a result. Notwithstanding, this work showed only very slight increases in accuracy (less than 1 %) as a result of pre-processing at these sites.

10.1.5 Image degradations combine to degrade results

The results show that applying downsampling, noise addition and compression operations to image sets result in derived products of worse quality than any of these operations independently. Images with compression of 'quality' 50 applied, noise of 21 db and downsampling to 81.5 % of their original size produced a median error of 7.62 mm for the Hunstanton results, with the downsampling operation alone producing an equivalent product with 7 mm error. Trade offs within image acquisition need to be properly rationalised and explained within scientific literature to ensure image quality is maximised for a given survey.

The recommendation from this is to keep in mind that optimising one of the parameters (compression, downsampling or noise) should not be done without considering the trade offs to the others.

10.1.6 Convergent, denser networks improve results

Where only limited numbers of images can be acquired or processed, images should be collected at narrowly oblique angles to one another, as results show that these image sets produced results competitive with the full image block. For example, set 3, with just 40 % of the images produced 90 % of the number of points of the full image set. The median error of set 3 was slightly higher at 6.65 mm as opposed to 6.36 mm for the full set. Whilst this has already been well studied in stereo contexts, and in simulations for mitigating systematic errors, the ordered sets of images within the study shows how value can be gained from intelligent network design.

The recommendation for this is to collect imagery at slightly convergent imaging angles (< 20° from the surface normal) to mitigate systematic error in the derived products. Densifying the camera network should be done

References

Agisoft (2018) *Agisoft Photoscan Professional Edition Last Accessed 27/08/2018.* Available at: http://www.agisoft.com/.

Ahmadabadian AH, Robson S, Boehm J, Shortis M, Wenzel K and Fritsch D (2013) A comparison of dense matching algorithms for scaled surface reconstruction using stereo camera rigs. *ISPRS Journal of Photogrammetry and Remote Sensing* 78: 157-167.

Barnhart, T.B. and Crosby, B.T., 2013. Comparing two methods of surface change detection on an evolving thermokarst using high-temporal-frequency terrestrial laser scanning, Selawik River, Alaska. *Remote Sensing*, *5*(6), pp.2813-2837.

Barrow HG, Tenenbaum JM, Bolles RC and Wolf HC (1977) Parametric correspondence and chamfer matching: Two new techniques for image matching. *Proceedings of the 5th International Joint Conference on Artificial Intelligence* 2: 659-663.

Bauwens S, Fayolle A, Gourlet-Fleury S, Ndjele LM, Mengal C and Lejeune P (2016) Terrestrial photogrammetry: a non-destructive method for modelling irregularly shaped tropical tree trunks. *Methods in Ecology and Evolution* 8: 460-471.

Bay H, Ess A, Tuytelaars T and Van Gool L (2008) Speeded-up robust features (SURF). *Computer Vision and Image Understanding* 110(3): 346-359.

Bemis SP, Micklethwaite S, Turner D, James MR, Akciz S, Thiele ST, et al. (2014) Ground-based and UAV-Based photogrammetry: A multi-scale, high-resolution mapping tool for structural geology and paleoseismology. *Journal of Structural Geology* 69A(0): 163-178.

Benedetti L, Corsini M, Cignoni P, Callieri M and Scopigno R (2012) Color to gray conversions in the context of stereo matching algorithms. *Machine Vision and Applications* 23(2): 327-348.

Besl PJ and McKay ND (1992) A method for registration of 3-D shapes. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 14(2): 239-256.

Bettio F, Gobbetti E, Merella E and Pintus R (2013) Improving the digitization of shape and color of 3D artworks in a cluttered environment. *Digital Heritage International Congress* 1: 23-30.

References

BGS (2018a) *BGS Lexicon of Named Rock Units - Hunstanton Formation..* Available at: http://www.bgs.ac.uk/lexicon/lexicon.cfm?pub=HUCK. *Last Accessed 27/08/2018*

BGS (2018b) *BGS Lexicon of Named Rock Units - Carstone Formation.* Available at: http://www.bgs.ac.uk/lexicon/lexicon.cfm?pub=CA. *Last Accessed 27/08/2018.*

BGS (2016) *BGS Lexicon of Named Rock Units - Ferriby Chalk Formation.* Available at: http://www.bgs.ac.uk/lexicon/lexicon.cfm?pub=FYCK. *Last Accessed 27/08/2018.*

Brown DC (1966) Decentering distortion of lenses. *Photometric Engineering* 32(3): 444-462.

Burns J, Delparte D, Gates R and Takabayashi M (2015) Integrating structure-from-motion photogrammetry with geospatial software as a novel technique for quantifying 3D ecological characteristics of coral reefs. *Peerj* 3: e1077.

Caduff R, Schlunegger F, Kos A and Wiesmann A (2015) A review of terrestrial radar interferometry for measuring surface change in the geosciences. *Earth Surface Processes and Landforms* 40(2): 208-228.

Canny J (1986) A computational approach to edge detection. *IEEE Transactions on Pattern Analysis and Machine Intelligence*(6): 679-698.

Castillo C, Pérez R, James M, Quinton J, Taguas EV and Gómez JA (2012) Comparing the accuracy of several field methods for measuring gully erosion. *Soil Science Society of America Journal* 76(4): 1319-1332.

Cerdan O, Govers G, Le Bissonnais Y, Van Oost K, Poesen J, Saby N, et al. (2010) Rates and spatial variations of soil erosion in Europe: a study based on erosion plot data. *Geomorphology* 122(1): 167-177.

Chen T, Catrysse PB, El Gamal A and Wandell BA (2000) How small should pixel size be? *Sensors and Camera Systems for Scientific, Industrial, and Digital Photography Applications* 3965: 451-459.

Chianucci F, Disperati L, Guzzi D, Bianchini D, Nardino V, Lastri C, et al. (2016) Estimation of canopy attributes in beech forests using true colour digital images from a small fixed-wing UAV. *International Journal of Applied Earth Observation and Geoinformation* 47: 60-68.

References

CloudCompare (2018) *EDF R&D Telecom Paris..* Available at: http://www.cloudcompare.org/. *Last Accessed 23/09/2018*

Cracknell MJ and Reading AM (2014) Geological mapping using remote sensing data: A comparison of five machine learning algorithms, their response to variations in the spatial distribution of training data and the use of explicit spatial information. *Computers & Geosciences* 63: 22-33.

Cronk S, Fraser C and Hanley H (2006) Automated metric calibration of colour digital cameras. *The Photogrammetric Record* 21(116): 355-372.

Cunliffe AM, Brazier RE and Anderson K (2016) Ultra-fine grain landscape-scale quantification of dryland vegetation structure with drone-acquired structure-from-motion photogrammetry. *Remote Sensing of Environment* 183: 129-143.

D'Amato J, Hantz D, Guerin A, Jaboyedoff M, Baillet L and Mariscal A (2016) Influence of meteorological factors on rockfall occurrence in a middle mountain limestone cliff. *Natural Hazards and Earth System Sciences* 16(3): 719-735.

DC-Raw (2018) *Decoding Raw Digital Photos in Linux* Available at: https://www.cybercom.net/~dcoffin/dcraw/. *Last Accessed 27/08/2018.*

Debevec PE and Malik J (1997) Recovering High Dynamic Range Radiance Maps from Photographs. *Siggraph*: 369-378.

DEFRA (2018) *Overstrand Cliffs - Special Area of Conservation - SAC - Habitats Directive.* Available at: http://jncc.defra.gov.uk/protectedsites/sacselection/sac.asp?EUCode=UK0030232. *Last Accessed 27/08/2018.*

Dold J (1996) Influence of large targets on the results of photogrammetric bundle adjustment. *International Archives of Photogrammetry and Remote Sensing* 31(B5): 119-123.

DxOMark (2018) *DxOMark by DxO* Available at: www.dxomark.com. *Last Accessed 27/08/2018.*

Eltner A, Baumgart P, Maas H and Faust D (2015) Multi-temporal UAV data for automatic measurement of rill and interrill erosion on loess soil. *Earth Surface Processes and Landforms* 40(6): 741-755.

References

Eltner A, Kaiser A, Abellan A and Schindewolf M (2017) Time lapse structure-from-motion photogrammetry for continuous geomorphic monitoring. *Earth Surface Processes and Landforms* 42(14): 2240-2253.

Eltner A, Mulsow C and Maas H (2013) Quantitative measurement of soil erosion from TLS and UAV data. *ISPRS-International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences* 1(2): 119-124.

Ely JC, Graham C, Barr ID, Rea BR, Spagnolo M and Evans J (2016) Using UAV acquired photography and structure from motion techniques for studying glacier landforms: application to the glacial flutes at Isfallsglaciären. *Earth Surface Processes and Landforms*.

Fan Z and De Queiroz RL (2003) Identification of bitmap compression history: JPEG detection and quantizer estimation. *IEEE Transactions on Image Processing* 12(2): 230-235.

Faro (2015) *Faro Edge Technical Sheet* Available at: www.faro.app.box.com/s/5p3kzuzqllhu2e5ea19mvd39pdbvooga/file/43487601589;. *Last Accessed 27/08/2018.*

Fischler MA and Bolles RC (1981) Random sample consensus: a paradigm for model fitting with applications to image analysis and automated cartography. *Communications of the ACM* 24(6): 381-395.

Fleischmann, P. Kneip, J. and Berns, K. (2016) An adaptive detection approach for autonomous forest path following using stereo vision. *International Conference on Control, Automation, Robotics and Vision*: 1-6.

Fonstad MA, Dietrich JT, Courville BC, Jensen JL and Carbonneau PE (2013) Topographic structure from motion: a new development in photogrammetric measurement. *Earth Surface Processes and Landforms* 38(4): 421-430.

Fraser CS (2013) Automatic camera calibration in close range photogrammetry. *Photogrammetric Engineering & Remote Sensing* 79(4): 381-388.

Fraser CS (1984) Network design considerations for non-topographic photogrammetry. *Photogrammetric Engineering and Remote Sensing* 50(8): 1115-1126.

References

Fraser CS (1996) Network design. In: Anonymous *Close Range Photogrammetry and Machine Vision*. Whittle Publising Bristol, 256-281.

FujiFilm (2018) *Fujifilm X-Trans Sensor Technology* Available at: https://www.fujifilm.eu/uk/products/digital-cameras/model/x-pro1/features-4483/aps-c-16m-x-trans-cmos. *Last Accessed 27/08/2018.*

Furukawa Y and Ponce J (2010) Accurate, dense, and robust multiview stereopsis. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 32(8): 1362-1376.

Genchi SA, Vitale AJ, Perillo GM and Delrieux CA (2015) Structure-from-motion approach for characterization of bioerosion patterns using UAV imagery. *Sensors* 15(2): 3593-3609.

Gigli G, Morelli S, Fornera S and Casagli N (2014) Terrestrial laser scanner and geomechanical surveys for the rapid evaluation of rock fall susceptibility scenarios. *Landslides* 11(1): 1-14.

Gómez-Gutiérrez, Á., de Sanjosé-Blasco, J. J., Lozano-Parra, J., Berenguer-Sempere, F., & de Matías-Bejarano, J. (2015). Does HDR pre-processing improve the accuracy of 3D models obtained by means of two conventional SfM-MVS software packages? The case of the Corral del Veleta rock glacier. *Remote sensing*, *7*(8), 10269-10294.

Gong P, Wang J, Yu L, Zhao Y, Zhao Y, Liang L, et al. (2013) Finer resolution observation and monitoring of global land cover:First mapping results with Landsat TM and ETM data. *International Journal of Remote Sensing* 34(7): 2607-2654.

Grafarend EW (1974) Optimization of Geodetic Networks. *Bolletino Di Geodesia a Science Affini* 33(4): 351-406.

GRASS GIS (2017) *Geographic Resources Analysis Support System (GRASS GIS) Software, Version 7.2.* Available at: http://grass.osgeo.org.

Grigillo D, Kosmatin Fras M and Petrovič D (2012) Automated building extraction from IKONOS images in suburban areas. *International Journal of Remote Sensing* 33(16): 5149-5170.

Grinand C, Rakotomalala F, Gond V, Vaudry R, Bernoux M and Vieilledent G (2013) Estimating deforestation in tropical humid and dry forests in Madagascar from 2000 to 2010 using multi-date Landsat satellite images and the random forests classifier. *Remote Sensing of Environment* 139: 68-80.

References

Gruen A (2012) Development and status of image matching in photogrammetry. *The Photogrammetric Record* 27(137): 36-57.

Grundland M and Dodgson NA (2007) Decolorize: Fast, contrast enhancing, color to grayscale conversion. *Pattern Recognition* 40(11): 2891-2896.

Guo R, Dai Q and Hoiem D (2011) Single-image shadow detection and removal using paired regions. *IEEE Computer Society Conference on Computer Vision and Pattern Recognition*: 2033-2040.

Hamilton E (1992) *JPEG File Interchange Format* Available at: https://www.repository.cam.ac.uk/bitstream/handle/1810/54/jfif.html?sequence=2. *Last Accessed 27/08/2018.*

Hanel ML, Kuhn S, Henrich D, Grune L and Pannek J (2012) Optimal camera placement to measure distances regarding static and dynamic obstacles. *International Journal of Sensor Networks* 12(1): 25-36.

Hänsel P, Schindewolf M, Eltner A, Kaiser A and Schmidt J (2016) Feasibility of High-Resolution Soil Erosion Measurements by Means of Rainfall Simulations and SfM Photogrammetry. *Hydrology* 3(4): 38.

Hasinoff SW (2014) Photon, Poisson Noise. In: Anonymous *Computer Vision.* : Springer, 608-610.

Hirschmüller H (2008) Stereo processing by semiglobal matching and mutual information. *Pattern Analysis and Machine Intelligence, IEEE Transactions on* 30(2): 328-341.

Hirschmüller H and Scharstein D (2009) Evaluation of stereo matching costs on images with radiometric differences. *Pattern Analysis and Machine Intelligence, IEEE Transactions on* 31(9): 1582-1599.

Homer CG, Dewitz JA, Yang L, Jin S, Danielson P, Xian G, et al. (2015) Completion of the 2011 National Land Cover Database for the conterminous United States-Representing a decade of land cover change information. *Photogrammetric Engineering and Remote Sensing* 81(5): 345-354.

Hytti HT (2006) Characterization of digital image noise properties based on RAW data. *Image Quality and System Performance III, Proceedings of SPIE-IS&T Electronic Imaging* 6059: 60590A-6059-12.

References

ImageMagick (2018) *Convert, Edit Or Compose Bitmap Images @ ImageMagick* Available at: http://www.imagemagick.org/script/index.php. *Last Accessed 23/09/2018.*

Immerzeel W, Kraaijenbrink P, Shea J, Shrestha A, Pellicciotti F, Bierkens M, et al. (2014) High-resolution monitoring of Himalayan glacier dynamics using unmanned aerial vehicles. *Remote Sensing of Environment* 150: 93-103.

International Organization for Standardization (2006) *12232: Photography-Electronic Still Picture Cameras: Determination of ISO Speed* Available at: https://www.iso.org/standard/37777.html. *Last Accessed 27/08/2018.*

ITU (1995) 601-5, Studio Encoding Parameters of Digital Television for Standard 4: 3 and Widescreen 16: 9 Aspect Ratios. *International Telecommunication Union* 96.

James MR, Ilic S and Ruzic I (2013) Measuring 3D coastal change with a digital camera. *Proceedings of the 7th International Conference on Coastal Dynamics*: 24-28.

James MR and Robson S (2012) Straightforward reconstruction of 3D surfaces and topography with a camera: Accuracy and geoscience application. *Journal of Geophysical Research: Earth Surface* 117.

James MR and Robson S (2014) Mitigating systematic error in topographic models derived from UAV and ground-based image networks. *Earth Surface Processes and Landforms* 39(10): 1413-1420.

James MR, Robson S, d'Oleire-Oltmanns S and Niethammer U (2017a) Optimising UAV topographic surveys processed with structure-from-motion: Ground control quality, quantity and bundle adjustment. *Geomorphology* 280: 51-66.

James, M.R., Robson, S. and Smith, M.W., (2017b). 3-D uncertainty-based topographic change detection with structure-from-motion photogrammetry: precision maps for ground control and directly georeferenced surveys. *Earth Surface Processes and Landforms*, *42*(12), pp.1769-1788.

Javernick L, Brasington J and Caruso B (2014) Modeling the topography of shallow braided rivers using Structure-from-Motion photogrammetry. *Geomorphology* 213(0): 166-182.

Johnson K, Nissen E, Saripalli S, Arrowsmith JR, McGarey P, Scharer K, et al. (2014) Rapid mapping of ultrafine fault zone topography with structure from motion. *Geosphere* 10(5): 969-986.

References

Kaiser A (2016) *Digital Elevation Models in 10 Minute Time Steps – a Status Report on 4D Monitoring of an Active Erosional Scar.* EGU PICO: EGU.

Kaiser A, Neugirg F, Rock G, Müller C, Haas F, Ries J, et al. (2014) Small-scale surface reconstruction and volume calculation of soil erosion in complex Moroccan gully morphology using structure from motion. *Remote Sensing* 6(8): 7050-7080.

Kanan C and Cottrell GW (2012) Color-to-grayscale: does the method matter in image recognition? *Plos One* 7(1): e29740.

Konolige K and Garage W (2010) Sparse Sparse Bundle Adjustment. *Proceedings of the British Machine Vision Conference* 10: 102.1-102.11.

Kontogianni, G., Chliverou, R., Koutsoudis, A., Pavlidis, G. and Georgopoulos, A., 2017. Enhancing Close-up image based 3D digitisation with focus stacking. *International Archives of the Photogrammetry, Remote Sensing & Spatial Information Sciences*, *42*.

Lague D, Brodu N and Leroux J (2013) Accurate 3D comparison of complex topography with terrestrial laser scanner: Application to the Rangitikei canyon (NZ). *ISPRS Journal of Photogrammetry and Remote Sensing* 82: 10-26.

Lane S, James T and Crowell M (2000) Application of digital photogrammetry to complex topography for geomorphological research. *The Photogrammetric Record* 16(95): 793-821.

Leica (2016) *Leica Monochrom m Technical Sheet.* Available at: http://uk.leica-camera.com/content/download/124054/1363404/version/5/file/Datenblatt_M+Monochrom+%28Typ+246%29_e.pdf. *Last Accessed 27/08/2018.*

Leica (2017a) *Leica P40 Datasheet* Available at: http://leica-geosystems.com/-/media/files/leicageosystems/products/datasheets/leica_scanstation_p30-p40_plant_ds.ashx?la=en. *Last Accessed 27/08/2018.*

Leica (2017b) *TCR 805 Manual* Available at: http://www.engineeringsurveyor.com/software/leica/leica_tc605_tc805_tc905.pdf. *Last Accessed 27/08/2018.*

LensScore (2018) *Digital Camera LensRating* Available at: http://www.lenscore.org/;. *Last Accessed 27/08/2018.*

References

Leprince S, Ayoub F, Klinger Y and Avouac J (2007) Co-registration of optically sensed images and correlation (COSI-Corr): An operational methodology for ground deformation measurements. *Geoscience and Remote Sensing Symposium*: 1943-1946.

Lillesand T, Kiefer RW and Chipman J (2014) *Remote Sensing and Image Interpretation.* : John Wiley & Sons.

Lisein J, Pierrot-Deseilligny M, Bonnet S and Lejeune P (2013) A photogrammetric workflow for the creation of a forest canopy height model from small unmanned aerial system imagery. *Forests* 4(4): 922-944.

Lourakis MI and Argyros AA (2009) SBA: A software package for generic sparse bundle adjustment. *ACM Transactions on Mathematical Software (TOMS)* 36(1): 2.

Lowe DG (2004) Distinctive image features from scale-invariant keypoints. *International Journal of Computer Vision* 60(2): 91-110.

Lowe DG (1999) Object recognition from local scale-invariant features. *Proceedings of the Seventh IEEE International Conference on Computer Vision* 2: 1150-1157.

Lu C, Xu L and Jia J (2012) Real-time contrast preserving decolorization. *SIGGRAPH Asia Technical Briefs*: 34-38.

Lucieer A, Jong SM and Turner D (2014) Mapping landslide displacements using Structure from Motion (SfM) and image correlation of multi-temporal UAV photography. *Progress in Physical Geography* 38: 97-116.

Lyon RF (2006) A brief history of'pixel'. *SPIE Symposium on Electronic Imaging*: 606901-606901-15.

MacDonald L, Toschi I, Nocerino E, Hess M, Remondino F and Robson S (2016) Accuracy Of 3D reconstruction in an illumination dome. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences* 41: 69-76.

Maini R and Aggarwal H (2010) A comprehensive review of image enhancement techniques. *Journal of Computing* 2(3): 8-13.

References

Margono BA, Turubanova S, Zhuravleva I, Potapov P, Tyukavina A, Baccini A, et al. (2012) Mapping and monitoring deforestation and forest degradation in Sumatra (Indonesia) using Landsat time series data sets from 1990 to 2010. *Environmental Research Letters* 7(3): 034010.

Mason S (1995) Expert system-based design of close-range photogrammetric networks. *ISPRS Journal of Photogrammetry and Remote Sensing* 50(5): 13-24.

Micheletti N, Chandler JH and Lane SN (2015) Structure from motion (SFM) photogrammetry. In: Anonymous *Geomorphological Techniques.* London, UK: British Society of Geomorphology.

Micklethwaite, S., Turner, D., Vasuki, Y., Kovesi, P. and Lucieer, A., 2012. Mapping from an Armchair: Rapid, high-resolution mapping using UAV and computer vision technology. *Proceedings of Structural Geology and Resources*, pp.130-133.

MicMac (2018) *Micmac – SOFTWARE.* Available at: http://logiciels.ign.fr/?Micmac. *Last Accessed 23/09/2018.*

Middlebury (2015) *Multi-View Stereo Evaluation Page.* Available at: http://vision.middlebury.edu/mview/eval/. *Last Accessed 27/08/2018.*

Mikita T, Janata P and Surový P (2016) Forest stand inventory based on combined aerial and terrestrial close-range photogrammetry. *Forests* 7(8): 165.

Mittal A, Moorthy AK and Bovik AC (2012) No-reference image quality assessment in the spatial domain. *IEEE Transactions on Image Processing* 21(12): 4695-4708.

Mlambo R, Woodhouse IH, Gerard F and Anderson K (2017) Structure from Motion (SfM) photogrammetry with drone data: a low cost method for monitoring greenhouse gas emissions from forests in developing countries. *Forests* 8(3): 68.

Mosbrucker AR, Major JJ, Spicer KR and Pitlick J (2017) Camera system considerations for geomorphic applications of SfM photogrammetry. *Earth Surface Processes and Landforms* 42(6): 969-986.

Nakano T, Kamiya I, Tobita M, Iwahashi J and Nakajima H (2014) Landform monitoring in active volcano by UAV and SfM-MVS technique. *ISPRS-International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences* 1: 71-75.

References

Niethammer U, Rothmund S, James M, Travelletti J and Joswig M (2010) UAV-based remote sensing of landslides. *International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences* 38(Part 5): 496-501.

Nikon (2017) *Nikon D700 Manual.* Available at: http://cdn-10.nikon-cdn.com/pdf/manuals/dslr/D700_en.pdf. *Last Accessed 27/08/2018.*

Nocerino E, Menna F and Remondino F (2014) Accuracy of typical photogrammetric networks in cultural heritage 3D modelling projects. *International Archives of the Photogrammetry, Remote Sensing & Spatial Information Sciences* 45.

Nouwakpo SK, James MR, Weltz MA, Huang C, Chagas I and Lima L (2014) Evaluation of structure from motion for soil microtopography measurement. *The Photogrammetric Record* 29(147): 297-316.

O'Connor J, Smith MJ and James MR (2016) Image inputs in Structure-from-Motion Photogrammetry: optimising image greyscaling. *EGU AGM* 18: 14902.

O'Connor J, Smith MJ and James MR (2017) Cameras and settings for aerial surveys in the geosciences: Optimising image data. *Progress in Physical Geography: Earth and Environment* 41(3): 325-344.

Olague G (2001) Autonomous Photogrammetric Network Design Using Genetic Algorithms. *Applications of Evolutionary Computing* 2037: 353-363.

OSGeo (2014) *Orthorectification of Images.* Available at: https://trac.osgeo.org/ossim/wiki/orthorectification. *Last Accessed 27/08/2018.*

Ouedraogo MM, Degra A, Debouche C and Lisein J (2014) The evaluation of unmanned aerial system-based photogrammetry and terrestrial laser scanning to generate DEMs of agricultural watersheds. *Geomorphology* 214(0): 339-355.

Paton, M, MacTavish, K, Ostafew, C.J. and Barfoot, T.D. (2015) It's not easy seeing green: Lighting-resistant stereo Visual Teach & Repeat using color-constant images. *2015 IEEE International Conference on Robotics and Automation (ICRA)*: 1519-1526.

Pix4D (2018) *Pix4D.* Available at: https://pix4d.com/. *Last Accessed 23/09/2018.*

References

Prosdocimi M, Burguet M, Di Prima S, Sofia G, Terol E, Comino JR, et al. (2017) Rainfall simulation and Structure-from-Motion photogrammetry for the analysis of soil water erosion in Mediterranean vineyards. *Science of the Total Environment* 574: 204-215.

Puliti S, Ørka HO, Gobakken T and Næsset E (2015) Inventory of small forest areas using an unmanned aerial system. *Remote Sensing* 7(8): 9632-9654.

Rasche K, Geist R and Westall J (2005) Detail preserving reproduction of color images for monochromats and dichromats. *Computer Graphics and Applications, IEEE* 25(3): 22-30.

Rieke-Zapp D and Nearing MA (2005) Digital close range photogrammetry for measurement of soil erosion. *Photogrammetric Record* 20: 69-87.

Rublee E, Rabaud V, Konolige K and Bradski G (2011) ORB: an efficient alternative to SIFT or SURF. *IEEE Computer Society Conference on Computer Vision and Pattern Recognition*: 2564-2571.

Ryan J, Hubbard A, Box J, Todd J, Christoffersen P, Carr J, et al. (2015) UAV photogrammetry and structure from motion to assess calving dynamics at Store Glacier, a large outlet draining the Greenland ice sheet. *The Cryosphere* 9(1): 1-11.

Sanz-Ablanedo E, Rodríguez-Pérez JR, Arias-Sánchez P and Armesto J (2009) Metric Potential of a 3D Measurement System Based on Digital Compact Cameras. *Sensors (Basel, Switzerland)* 9(6): 4178-4194.

Shortis MR and Seager JW (2014) A practical target recognition system for close range photogrammetry. *The Photogrammetric Record* 29(147): 337-355.

Sieberth T, Wackrow R and Chandler JH (2013) Automatic isolation of blurred images from UAV image sequences. *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences* XL-1/W2: 4-6.

Sieberth T, Wackrow R and Chandler JH (2014) Motion blur disturbs - the influence of motion-blurred images in photogrammetry. *The Photogrammetric Record* 29(148): 434-453.

Smith MW and Vericat D (2015) From experimental plots to experimental landscapes: topography, erosion and deposition in sub-humid badlands from Structure-from-Motion photogrammetry. *Earth Surface Processes and Landforms* 40(12): 1656-1671.

References

Snavely N, Seitz SM and Szeliski R (2008) Modeling the world from internet photo collections. *International Journal of Computer Vision* 80(2): 189-210.

Strecha C, von Hansen W, Gool LV, Fua P and Thoennessen U (2008) On benchmarking camera calibration and multi-view stereo for high resolution imagery. *IEEE Conference on Computer Vision and Pattern Recognition*: 1-8. Anchorage, AK, USA

Stumpf A, Malet J, Allemand P, Pierrot-Deseilligny M and Skupinski G (2015) Ground-based multi-view photogrammetry for the monitoring of landslide deformation and erosion. *Geomorphology* 231: 130-145.

Tavani S, Corradetti A and Billi A (2016) High precision analysis of an embryonic extensional fault-related fold using 3D orthorectified virtual outcrops: The viewpoint importance in structural geology. *Journal of Structural Geology* 86: 200-210.

Tipler PA (2004) *Physics for Scientists and Engineers.* New York: New York : W. H. Freeman, pp 1091 – 1093.

Triggs B, McLauchlan PF, Hartley RI and Fitzgibbon AW (1999) Bundle adjustment—a modern synthesis. *International Workshop on Vision Algorithms*: 298-372.

Turner D, Lucieer A and de Jong SM (2015) Time series analysis of landslide dynamics using an unmanned aerial vehicle (UAV). *Remote Sensing* 7(2): 1736-1757.

UK Environment Agency (2018) *Environment Agency Geomatics Survey Data.* Available at: https://environment.data.gov.uk/ds/survey/. *Last Accessed 27/08/2018.*

Van der Meer, Freek D, Van der Werff, Harald MA, van Ruitenbeek FJ, Hecker CA, Bakker WH, Noomen MF, et al. (2012) Multi-and hyperspectral geologic remote sensing: A review. *International Journal of Applied Earth Observation and Geoinformation* 14(1): 112-128.

Vasuki Y, Holden E, Kovesi P and Micklethwaite S (2014) Semi-automatic mapping of geological Structures using UAV-based photogrammetric data: An image analysis approach. *Computers & Geosciences* 69: 22-32.

Verhoeven G (2010) It's all about the format–unleashing the power of RAW aerial photography. *International Journal of Remote Sensing* 31(8): 2009-2042.

References

Verhoeven G, Karel W, Štuhec S, Doneus M, Trinks I and Pfeifer N (2015) Mind your grey tones–examining the influence of decolourization methods on interest point extraction and matching for architectural image-based modelling. *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, XL* XL-5/W4: 307-314.

VisualSFM (2018) *VisualSFM : A Visual Structure from Motion System.* Available at: http://ccwu.me/vsfm/. *Last Accessed 23/09/2018.*

Wackrow R and Chandler JH (2011) Minimising systematic error surfaces in digital elevation models using oblique convergent imagery. *The Photogrammetric Record* 26(133): 16-31.

Wallace L, Lucieer A, Malenovský Z, Turner D and Vopěnka P (2016) Assessment of forest structure using two UAV techniques: A comparison of airborne laser scanning and structure from motion (SfM) point clouds. *Forests* 7(3): 62.

Wang Z, Sheikh HR and Bovik AC (2002) No-reference perceptual quality assessment of JPEG compressed images. *Proceedings. 2002 International Conference on Image Processing* 1: I-I.

Westoby M, Brasington J, Glasser N, Hambrey M and Reynolds J (2012) 'Structure-from-Motion'photogrammetry: A low-cost, effective tool for geoscience applications. *Geomorphology* 179: 300-314.

Whitehead K, Moorman B and Hugenholtz C (2013) Brief Communication: Low-cost, on-demand aerial photogrammetry for glaciological measurement. *The Cryosphere* 7(6): 1879-1884.

Williams R, Brasington J, Vericat D and Hicks D (2014) Hyperscale terrain modelling of braided rivers: fusing mobile terrestrial laser scanning and optical bathymetric mapping. *Earth Surface Processes and Landforms* 39(2): 167-183.

Wood C (1967) Some new observations on the Maestrichtian Stage in the British Isles. *Bulletin of the Geological Survey of Great Britain* 27: 271-288.

Woodget, A.S., Visser, F., Maddock, I.P., Carbonneau, P. and Austrums, R., 2016. Quantifying fluvial substrate size using hyperspatial resolution UAS imagery and SfM-photogrammetry. In *Extended Abstract, 11th International Symposium on Ecohydraulics*.

Wu C (2014) Critical configurations for radial distortion self-calibration. *IEEE Conference on Computer Vision and Pattern Recognition*: 25-32.

References

Zbontar J and LeCun Y (2016) Stereo Matching by Training a Convolutional Neural Network to Compare Image Patches. *Journal of Machine Learning Research* 17(65): 1-32.

Zhang B, Miller S, Walker S and Devenecia K (2007) Next generation automatic terrain extraction using Microsoft UltraCam imagery. *Proceedings of the ASPS Annual Conference*.

Zhang Z (2000) A flexible new technique for camera calibration. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 22(11): 1330-1334.

Zielke O, Arrowsmith JR, Grant Ludwig L and Akciz SO (2010) Slip in the 1857 and earlier large earthquakes along the Carrizo Plain, San Andreas Fault. *Science* 327(5969): 1119-1122.

Appendix

```
import PhotoScan


doc = PhotoScan.app.document
chunks = doc.chunks
for chunk in chunks:
   frame = chunk.frames[0]
   frame.matchPhotos(accuracy=PhotoScan.HighestAccuracy, keypoint_limit=10000000,
tiepoint_limit=10000000)
   chunk.alignCameras()
   chunk.optimizeCameras(fit_b1=False, fit_b2=False, fit_k2=False, fit_k3=False, fit_k4=False,
fit_p1=False, fit_p2=False)
   chunk.buildDenseCloud(quality=PhotoScan.UltraQuality)
```

Appendix 1. Script used for the standard workflow within Photoscan. This was executed **after** loading each set of images into separate chunks.

Contrast optimization for structure-from-motion surveys

James O'Connor[1] (James.OConnor@kingston.ac.uk), Mike Smith[1], Mike R. James[2]
[1]School of the Natural and Built Environment, Kingston University, [2]Lancaster Environment Centre, Lancaster University

Environment Centre | Lancaster University | Kingston University London

1. Introduction

Motivation: Structure-from-motion (SfM) is a technology which takes multiple 2D photos of a common scene as an input and projects the data into 3D. Feature based SfM pipelines take single channel (grayscale) images as inputs. Within this contribution we review the results from varying this channel.

Research question: Can we improve the accuracy of SfM models by generating appropriate greyscale channels from RGB images?

2. Methods

a. SfM software takes *greyscale* images as inputs

b. By altering this greyscale input, we intend to optimise contrast and improve outputs

c. **Experiment:** Take greyscale images made using different combinations of the Red, Green and Blue colour bands. Test other operators and compare accuracies

d. Measure **accuracy** against a very high quality terrestrial laser scan

e. **Create** greyscale image subsets generated by imposing the constraint that the Red, Green and Blue image bands must sum to 1. Bands are incremented by .05 per set and all permutations generated

f. 231 image subsets are constructed with 10 models generated for each set. Error (Cloud-to-truth) surfaces are then generated for the entire dataset

g. Data **augmented** with further algorithms

h. Two datasets used, one a 9 image block of a Welsh coastal cliff (Figure 2, Westoby et. al 2012) and one a 21 image block of a lab controlled scene (courtesy of Annette Eltner, figure 2)

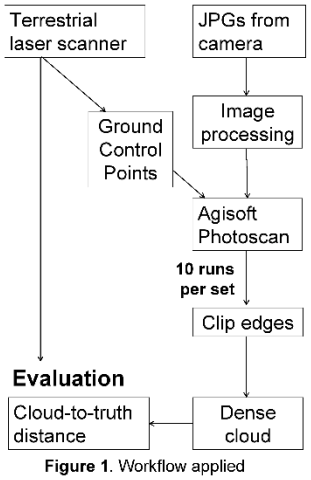

Figure 1. Workflow applied

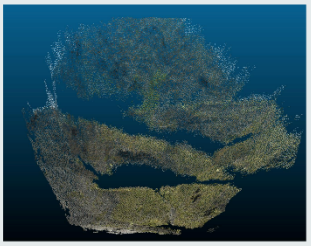

Figure 2. Constitution Hill overview

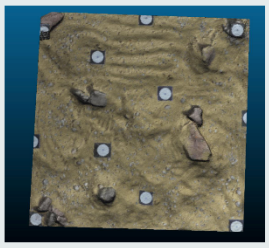

Figure 3. Lab control overview

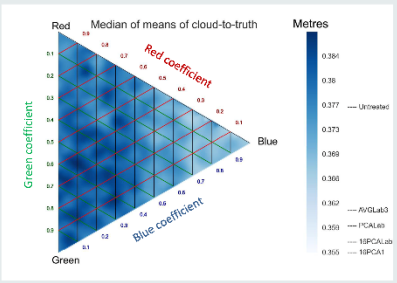

Figure 4. Results from Constitution Hill block

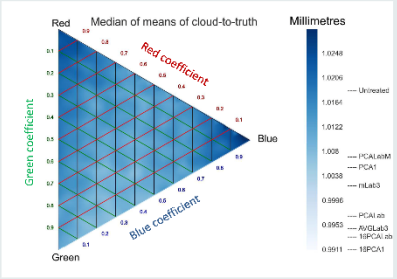

Figure 5. Results from Lab control block

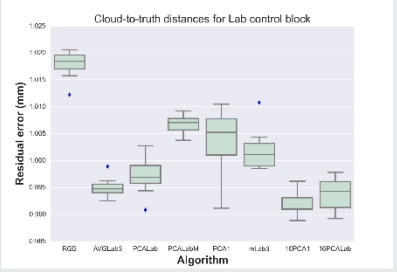

Figure 6. Results of other operations for lab control block

3. Description of other algorithms

PCA1 : First Principal Component (PC) of each RGB image

PCALab : First PC of each image after conversion to LAB colourspace

16PCA1/16PCALab: 16 bit First PC of first two algorithms

AVGLab3 : Taken from Verhoeven et al. (2015), seeks to maximise statistical information in a single image

PCALabM : Multi-image decolourization. Images are concatenated before taking first PC in order to ensure global consistency (Bennedetti et al. 2012)

mLab3 : AVGLab3, but with masking of unimportant regions before applying the workflow

4. Results

- Combinations of RGB images mapped onto a triangle, showing distance to reference laser scan (Figure 4 and 5)
- Median of mean absolute cloud-to-truth distance for 10 runs presented
- Other algorithms listed in colourbar and figure 6

5. Discussion

- Cloud to cloud errors shown to be reduced by up to 6% (Constitution hill) vs untreated images using simple contrast enhancement techniques
- Multi-image decolorization shown to not perform as well as image-specific techniques in dataset b
- Range of differences appear to depend on number of images within each block

6. Future work

- Bit depth will be investigated to see if further contrast enhancement can be achieved
- Number and geometry of images further studied to see if accuracy can be increased
- RAW image workflows currently being investigated to maximise value from consumer-grade cameras

Acknowledgments

Annette Eltner and Matt Westoby for providing the datasets. Geert Verhoever for sharing the code from his work

Contact: jp-oconnor.com

References

Benedetti L, Corsini M, Cignoni P, Callieri M and Scopigno R (2012) Color to gray conversions in the context of stereo matching algorithms. *Machine Vision and Applications* 23(2): 327-348

Verhoeven G, Karel W, Štuhec S, Doneus M, Trinks I and Pfeifer N (2015) *Mind Your Grey Tones—examining the Influence of Decolourization Methods on Interest Point Extraction and Matching for Architectural Image-Based Modelling*. : ISPRS

Westoby M, Brasington J, Glasser N, Hambrey M and Reynolds J (2012) 'Structure-from-Motion'photogrammetry: A low-cost, effective tool for geoscience applications. *Geomorphology* 179: 300-314.

Appendix 2. Poster presented at the European Geophysical Union's AGM, 2016.

```
import cv2
import numpy as np
import sklearn.decomposition as dc


def PCALab(filepath):
    """
    Python implementation of Verhoeven et al. (2015) avgLabPCAPCAwLabPCA
algorithm
    :param filepath: Path to TIF file
    :return: Processed image
    """
    image = cv2.imread(filepath)
    out_image = np.zeros_like(image).astype(float)
    lab_image = cv2.cvtColor(image, cv2.COLOR_RGB2Lab)
    pc = dc.pca.PCA(n_components=1)

    # Generate image 1
    lab_columns = np.reshape(lab_image, (lab_image.shape[0] *
lab_image.shape[1], 3))
    pc.fit(lab_columns)
    image_one = pc.transform(lab_columns)
    image_one = np.reshape(image_one, (lab_image.shape[0],
lab_image.shape[1]))
    image_one_out = (image_one - np.min(image_one)) / (np.max(image_one)
- np.min(image_one))

    for n in range(1, 4):
        # Generate weighted image data
        weighting = 0.25 + (n - 1) * 0.2
        band_coefficients = [1 - weighting, weighting / 2, weighting / 2]
        lab_columns = np.reshape(lab_image * band_coefficients, (1 -
lab_image.shape[0] * lab_image.shape[1], 3))

        # Perform PCA
        pc.fit(lab_columns)

        # Reshape data to image
        pc_data = pc.transform(lab_columns)
        pc_data = np.reshape(pc_data, (lab_image.shape[0],
lab_image.shape[1]))
        normed_data = (pc_data - np.min(pc_data)) / (np.max(pc_data) -
np.min(pc_data))
        out_image[:, :, n - 1] = normed_data

    three_band_pca_image = np.reshape(out_image, (out_image.shape[0] *
out_image.shape[1], 3))

    # Perform PCA on three band output
    pc.fit(three_band_pca_image)
    pca_pca_image = pc.transform(three_band_pca_image)
    pca_pca_image = np.reshape(pca_pca_image, (lab_image.shape[0],
lab_image.shape[1]))

    # Generate image 2
    normed_pca_pca_image = (pca_pca_image - np.min(pca_pca_image)) /
(np.max(pca_pca_image) - np.min(pca_pca_image))

    out_data = (normed_pca_pca_image + image_one_out) / 2
    return out_data
```

Appendix 3. Python function for generating PCALab images