Three-Dimensional Reconstruction of Braided River Morphology and Morphodynamics with Structure-from-Motion Photogrammetry

Joe Steven James

A thesis submitted in partial fulfilment of the requirements of the degree of Doctor of Philosophy.

School of Geography, Queen Mary, University of London

2017

Declaration and Statements

I, Joe Steven James, confirm that the research included within this thesis is my own work or that where it has been carried out in collaboration with, or supported by others, that this is duly acknowledged below and my contribution indicated. Previously published material is also acknowledged below.

I attest that I have exercised reasonable care to ensure that the work is original, and does not to the best of my knowledge break any UK law, infringe any third party's copyright or other Intellectual Property Right, or contain any confidential material.

I accept that the College has the right to use plagiarism detection software to check the electronic version of the thesis.

I confirm that this thesis has not been previously submitted for the award of a degree by this or any other university.

Signature:

Date: 20/09/2017

Details of collaborations:

The Dart research team, for data capture in both 2014 and 2015 under the NERC Urgent Grant (NE/M005054/1)

This work has been supported by a Natural Environmental Research Council studentship (Grant number NE/L501797/1

Acknowledgments

I would like to thank my primary supervisor Professor James Brasington, for the support, guidance and understanding throughout my thesis. Through James I have had the opportunity to travel, meet and work with a number of people who have enriched my PhD experience greatly. Without the chance to work on the Dart River, I can honestly say I may have never made it as far as New Zealand under my own steam and for that I am eternally grateful. Although having tasted the standard of his cooking from a fairly limited store cupboard, I now realise he is wasted in academia. I would also like to thank Dr Alex Henshaw who has provided guidance and a reassuring presence on a number of occasions, most notably before my first conference experience.

This thesis has required extensive fieldwork in New Zealand which would not have been possible without the following people; Simon Cook, Simon Cox, Eliisa Lotsari, Sam McColl, Richard Williams, and Damia Vericat. I am thankful for all of the help received and the chance to work alongside all of you. Most importantly, Niall Lehane has been a constant presence throughout, with work in both London and New Zealand, providing a much needed touch of optimism and sharing a passion for the combination of cookies and fieldwork.

I have had the privilege to work alongside a number of current and former PhD students through my time at QMUL who have all helped shape my experience. In particular I would like to thank Francis O'Shea for being my mentor when I first begun and Eleanor Webster for putting up with me. Beyond them, everyone I have met in Rooms 224 and 104 have genuinely helped this process run a lot smoother and I'm genuinely sad to be leaving Mile End.

Finally I'd like to thank my Mum, Dad and Sister for their support throughout and providing me with a sense of perspective.

Abstract

The recent emergence of Structure-from-Motion Photogrammetry (SfM) has created a cost-effective alternative to conventional laser scanning for the production of high-resolution topographic datasets. There has been an explosion of applications of SfM within the geomorphological community in recent years, however, the focus of these has largely been small-scale (10^2 – 10^3 m^2), building on innovations in low altitude Unmanned Aircraft Systems (UAS). This thesis examines the potential to extend the scope of SfM photogrammetry in order to quantify of landscape scale processes. This is examined through repeat surveys of a ~35 km^2 reach of the Dart River, New Zealand. An initial SfM survey of this reach was conducted in April 2014, following a large landslide at the Slipstream debris fan. Validation of the resulting digital elevation models using Independent Control Point's (ICPs) suggested encouraging results, however benchmarking the survey against a long-range laser scanned surface indicated the presence of significant systematic errors associated with inaccurate estimation of the SfM bundle adjustment. Using a combination of scaled laboratory field experiments, this research aimed to develop and test photogrammetric data collection and modelling strategies to enhance modelling of 3D scene structure using limited constraints. A repeat survey in 2015 provided an opportunity to evaluate a new survey strategy, incorporating a convergent camera network and a priori measurement of camera pose. This resulted in halving of mean checkpoint residuals and a reduction in systematic error. The models produced for both 2014 and 2015 were compared using a DEM differencing (DoD) methodology to assess the applicability of wide-area SfM models for the analysis of geomorphic change detection. The systematic errors within the 2014 model confound reliable change detection, although strategies to correlate the two surveys and measure the residual change show promise. The future use of SfM over broad landscape scales has significant potential, however, this will require robust data collection and modelling strategies and improved error modelling to increase user confidence.

Table of Contents

Glossary of key terms

AGL	Above Ground Level
APS	Agisoft PhotoScan
DEM	Digital Elevation Model
DoD	Dem-of-Difference
FOV	Field Of View
GCP	Ground Control Point
GPS	Geographical Positioning System
ICP	Independent Control Point
K1	Primary radial distortion parameter
LoD	Level of Detection
LOOCV	Leave One Out Cross Validation
MAE	Mean Absolute Error
MaxE	Maximum Error
ME	Mean Error
MVS	Multi-View Stereo
RMSE	Root Mean Square Error
RTK	Real Time Kinetic
SD	Standard Deviation
SDE	Standard Deviation of Error
SfM	Structure-from-Motion
SGM	Semi Global Matching
TLS	Terrestrial Laser Scanner
ToPCAT	Topographic Point Cloud Analysis Toolkit
UAS	Unmanned Aerial System

List of Figures

List of Tables

1. Chapter 1: Introduction

1.1. Shining a Light on Geomorphology: The Emergence of High Resolution Topography

The production of high-resolution topographic (HRT) datasets is of increasing interest within the geomorphological community (Fonstad et al., 2013). Historically, geomorphological studies have been limited by the available spatial and temporal resolution of data (Lane et al., 1994), however, the recent revolution in geomatics has transformed the quantification of landforms and landscapes from a data poor, to a data rich discipline (Brasington et al., 2012). Initially this was stimulated through the use of Total Stations and GNSS ground surveys, providing sub centimetre accuracy albeit at low point densities (Brasington et al., 2000; Westoby et al., 2012). Subsequently, the uptake of laser ranging techniques (deployed from both aerial and terrestrial platforms) has seen a dramatic step-change in survey point densities and spatial extents (Westoby et al., 2012; Smith et al., 2016). Airborne laser ranging (ALS), based on Light Detection And Ranging (LiDAR) is capable of capturing areas of up to 10^2 km^2 with 3D point accuracies in the order of ~10^{-1} m, enabling catchment-wide surveys (Notebaert et al., 2009). Terrestrial Laser Scanning (TLS), the ground based counterpart to ALS in contrast, is largely restricted to smaller spatial domains. However, over shorter ranges, the high frequency scanners used are able to capture data with sub-centimetre accuracy and with densities that exceed 10^3 points per m^{-2} (Wheaton et al., 2010; Williams et al., 2011; Brasington et al., 2012). Solutions to mount TLS on mobile platforms and support these with accompanying inertial navigation systems, now create the opportunity to cover greater survey areas, extending the application of these very high resolution survey tools (e.g., Vaaja et al., 2013).

1.2. Geomorphological Applications

The advances in topographic survey methods described above have revolutionized quantitative models of the Earth's surface morphology (Wheaton et al., 2010; Brasington et al., 2012). In a fluvial context, this has seen cross-section based models of channel geometry be progressively replaced by continuous 2.5D (z = f(x,y)) Digital Elevation Models, or DEMs, and more recently fully 3D point cloud datasets (Brasington et al., 2012; Javernick et al., 2014; Tarolli, 2014). With the capacity to capture 3D data at the resolution of individual particles and over extents that span not just bars and

channels, but growing towards entire catchments, there now exists the unparalleled prospect of analysing landforms and their organization from the scale of their basic building blocks upwards. This approach has the potential to transcend conventional scale boundaries that arguably reflect historic constraints on sampling design and observational methods rather than natural length scales of the environment. Indeed, Miller and Goodchild (2015) have suggested that the emergence of such geospatial 'Big Data' could herald a new approach, or as they put it, a fourth phase of science, across a wide range of spatial disciplines. They point to a future characterized by exploratory abduction of theory from this growing wealth of georeferenced data. While we have yet to see this approach become established in geomorphology, recent work by Redolfi et al. (2016) and Sangireddy et al. (2017) has demonstrated how the interrogation of dense terrain data can yield new insights into the structure of fluvial and catchment topography.

High Resolution Topography (HRT) is more commonly used to provide ever more detailed boundary conditions for numerical models and to develop novel parameterizations for conventionally sub-grid properties such as roughness and particle size (McMillan and Brasington, 2007; Brasington et al., 2012; Javernick et al., 2015). However, perhaps the most significant untapped opportunity arising from the emergence of HRT is the ability to quantify geomorphological change directly, accurately and contextually, through the comparison of digital terrain models through time. This approach is well-established methodologically (Vericat et al., 2017), but is rarely used to significant effect due to the difficulties of obtaining repeat surveys at frequencies tuned effectively to the key geophysical forcing processes. The use of repeat topographic models offers a truly 4-dimensional empirical analysis, enabling the development of sediment budgets that can account for longitudinal redistribution and storage of sediment within the channel and its floodplain. Beyond direct insight into morphological adjustment, or landscape morphodynamics, repeat topography also provides novel data to test the growing range of numerical simulation models that couple hydrodynamics and sediment transport to predict landscape evolution through time (Williams et al., 2016a).

While the rate of progress has been substantial, the continued use of laser ranging techniques to answer landscape-scale evolution questions remains limited by the substantial financial investment required in equipment, deployment and user training. This results in limited flexibility (or very high costs) of deployment and often poor manoeuvrability in difficult terrain (James and Robson, 2012). The result often therefore, is infrequent surveys, often commissioned only by major survey agencies, and at schedules that bear little reflection upon the forcing geophysical processes that shape landscapes (Brasington et al., 2012; Hugenholtz et al., 2013). Consequently, many 4D studies of landscapes are informed by serendipitous 'snapshots' of the topography, that often integrate multiple forcing events (storms, landslides, earthquakes), rendering direct analysis of cause and effect complex.

There is, therefore, a pressing requirement to develop cost-effective data acquisition techniques that offer comparable resolution and fidelity to laser ranging, but which enable frequent and flexible deployment to maintain adequate spatial and temporal coverage. In this context, the emergence and refinement of Structure-from-Motion (SfM) Photogrammetry holds considerable promise as a low-cost and largely automated method to address these issues (Westoby et al., 2012; Fonstad et al., 2013; Smith et al., 2016).

1.3. Low-cost Topography: Structure-from-Motion Photogrammetry

Structure from Motion Photogrammetry (hereafter, SfM for brevity) provides an inexpensive approach to reconstruct 3D point clouds that have comparable density and accuracy to laser ranging techniques and, moreover, may be applicable in contexts where complex terrain can preclude ALS or TLS (Brasington et al., 2012; James and Robson, 2012). Developments in this field stem from Marr and Poggio's (1979) seminal work on stereo vision, but have only recently been adopted beyond the computer vision community. This latest interest has given rise to the advent of a wide range of automated, consumer grade software packages that simplify the photogrammetric workflow into an automated 'pipeline' of processes that can be implemented with relatively little experience (Snavely et al., 2006; Verhoeven, 2011).

Unlike traditional softcopy photogrammetry, SfM relies on highly redundant, overlapping photographs which can be acquired from low-cost, non-metric and even uncalibrated compact cameras (James and Robson, 2012; Westoby et al., 2012; Fonstad et al., 2013). The techniques underpinning 3D reconstruction by SfM are reviewed in section 2.1. Briefly, in line with conventional approaches to photogrammetric reconstruction, the approach seeks to solve the intrinsic and external camera parameters to determine camera pose and use correspondence analysis across multiple images to locate projected 3D points (Cipolla, 2008). This emerging survey method, offers the opportunity to capture 3D data over wide spatial extents based on relatively cheap photography, whilst maintaining the high spatial resolution and fidelity associated with established laser ranging methods (Fonstad et al., 2013; Solbo and Storvold, 2013).

SfM Photogrammetry has been applied across a wide range of contexts, from architecture and archaeology, to computer gaming and mining and is ideally suited to topographic surveying (Luhmann et al., 20134). However, its application to specifically fluvial terrain modelling remains challenging due to subtle, shallow and morphological complex topography, vegetation cover and inundated channels. In addition, the high rates of topographic change experienced in labile rivers may also hinder establishing reliable, stable ground-control, particularly over large areas (Javernick et al., 2014).

1.4. Modelling the Morphology and Morphodynamics of Braided Rivers

Despite the challenges, there exists significant potential to apply photogrammetric methods in large (10^2-10^3 m wide), braided rivers. These rivers have traditionally been difficult to quantify using traditional survey methods, due to their high width, significant longitudinal and transverse morphological complexity and high rates of turnover. Nonetheless, considerable interest in the development of detailed 3D models of braided rivers has persisted over the last two decades. These data provide a means to estimate critical management information, including changes in channel capacity and bed load transport through DEM differencing (Brasington et al., 2003; Lane et al., 2003).

While characterized by typically low relief, braided systems are also associated with a high percentage of sub-aerial bed exposure at low flow (exceeding 85-90% of the channel width, e.g., Williams et al., 2013). The use of SfM in such a context is not without

complicating factors however. The majority of land survey applications of SfM and the related analysis workflows that have evolved elsewhere have largely been developed for relatively small spatial problems (0.1-1 km^2), often using photography acquired terrestrially, or from remotely piloted or Unmanned Aircraft Systems (UAS).

Extending these methods to cover not just wide braided rivers, but over sufficient longitudinal (10^{1-2} km), even catchment-scale, extents (10^2 km^2), in order to provide the appropriate contextualization of the geomorphological dynamics poses new challenges, logistically, technically and computationally. A key area for future research therefore, is the development of workflows for such broad-area reconstruction (Smith et al., 2016).

Additionally, and somewhat less commonly reported in the literature, the comparison of DEMs through time may also reveal the presence of complex systematic errors, such as datum shifts, navigation track bias and errors arising from block bundle adjustment that might otherwise go undetected in a single terrain model. The use of SfM to derive terrain models suitable for change detection and the quantification of channel morphodynamics, will therefore require a robust examination of DEM quality in order to distinguish signal from methodological error and ambiguities (Brasington et al., 2003; James et al., 2013).

1.5. A Geomorphological Context

Rivers play a vital role as the gutters of the landscape, conveying water and sediment from proximal hillslope sources to distal sedimentary sinks in lakes and oceans. Sediment delivery to rivers in Alpine environments is characterized by large but infrequent pulses of material sourced from landslides and debris flows. In extreme cases, when the rate of sediment supply exceeds the capacity of a channel to remove it, a landslide dam can form, impounding the flow and creating an inline lake. Such events are taken to be rare, although in tectonically active environments widespread co-seismic slope failures leading to landslide dams may be more common than anticipated. In the recent 7.8 Mw Kaikoura Earthquake in New Zealand in November 2016, an estimated 150 landslides severed channels creating in-line lakes (Massey et al., 2017).

Such events play a crucial but weakly understood role in the evolution of catchment drainage, channel morphology and sediment flux. Considerable attention has focused historically on the risks posed by the catastrophic failure of such landslide dams (Korup, 2005; Davies et al., 2007). However, in cases where landslides only partially or temporarily impede drainage, fluvial reworking of the sediment pulse results in a complex and less well-understood set of processes and hazards. These include cyclical channel aggradation and degradation that may destabilise floodplain assets, reduce standards of flood protection, elevate the risk of channel avulsion, and impact on freshwater and riparian ecology with a legacy that outlasts the initial disturbance (James, 2010).

Since the seminal work of Gilbert (1917), large pulses of sediment introduced into rivers have been hypothesized to travel downstream as translating bed waves, analogous to storm hydrographs. These disturbances are identifiable as transitory episodes of aggradation and degradation that propagate streamwise. The results of experimental and numerical modelling have questioned this conceptual model (Lisle et al., 2001; Cui et al., 2005) and suggest rather, that under the high Froude numbers typical of mountain rivers, sediment pulses are more likely to be dispersed symmetrically about the point of supply, with consequently more localised impacts. Field support for either model is currently limited and ambiguous, with observed disturbances found to dissipate through both translation (Pierson et al., 2011) and dispersion (Sutherland et al., 2002), or a combination of both (see James, 2010 for a review). Such uncertainty in the behaviour of sediment pulses urgently demands further research, especially as climate change has been linked to an intensification of landsliding in Alpine environments (Huggel et al., 2012).

Recent events on the Dart River, New Zealand (44.59 S 168.34 E) have provided an ideal opportunity to study the morphological evolution of a 35 km reach as it adjusts in the aftermath of a significant episode of allogenic sediment delivery. Lying on the eastern side of the Cosmos Peaks, the Slipstream (Māori: Te Horo) landslide covers an area of 0.9 km^2 with a steep toe that has contributed $10^5 – 10^6$ m^3 of sediment annually to the valley floor from 2009-2013 (Cox et al., 2014). The channel response to such a significant sediment source is unknown, a product of a lack of empirical studies and the complex nature of the catchment morphology.

Given the potential for sediment pulses to manifest as migratory episodes of aggradation and degradation, with associated impacts on channel capacity, avulsion risk and destabilization of riparian assets, there is an urgent need to quantify the baseline channel form, against which future changes in channel character can be assessed. Particular areas for concern include damage to the only bridge crossing at the Hillocks (44.773 S, 168.327 E), the progradation and destabilization of the Dart River delta into Lake Wakatipu and associated increased flood risk for local communities (ORC, 2008). Whilst the majority of modelling based studies suggest that such pulses are likely to disperse in situ (Lisle et al., 2001; Cao and Carling, 2003; Cui and Parker., 2003), previous empirical research has identified pulses migrating at rates of between 500-1600 m per anum, reaching apex heights of up to 23 m in confined regions/pinch points (Griffiths, 1993; Madej and Ozaki, 1996; Sutherland et al., 2002; Pierson et al., 2011). This is particularly relevant to the Dart River as the channel immediately downstream of the slide is confined within a relatively narrow gorge.

This background context was given recent impetus, when on 4/1/14 a major debris flow at Slip Stream introduced $>10^6$ m^3 of sediment to the Dart River valley floor. Runout over the existing fan dammed the Dart River causing a sudden drop in discharge downstream. This broad dam was breached quickly, however, the loss of conveyance impounded a 4 km long lake with a volume estimated at $>5 \times 10^6$ m^3 with depths exceeding 12 m. Research is therefore urgently needed to quantify both the initial state of the system in the aftermath of the major landslide in January 2014 and the subsequent morphological adjustment. Structure-from-Motion provides the opportunity to undertake such a catchment-scale study within practical financial constraints.

1.6. Thesis Aim and Objectives

The preceding sections have highlighted the potential of SfM as an emerging high resolution survey technique for deriving catchment scale data products. This thesis focuses upon the variation in the longitudinal sediment distribution between two epochs, requiring the construction of consistent DEMs to enable change detection analysis. As such the overarching aim of this thesis can be stated as:

- ❖ Assessing the utility of SfM Photogrammetry to derive catchment scale data products for the purposes of performing geomorphological change detection analysis.

This aim may further be divided into three distinct research goals:

A. To develop a field-to-data product workflow for 3D image-based topographic reconstruction of large fluvial systems;
B. To develop an error model to quantify the quality of SfM data products for modelling fluvial systems;
C. Apply (A and B) to assess the evolution of the Dart River from Slipstream to the Wakatipu delta between 2014 and 2015.

These will be achieved through by addressing a series of objectives:

- Conduct an initial SfM survey of the 35 km reach of the Dart River downstream of the Slipstream landslide complex to the delta using conventional methods and critically evaluate the resulting quality of the terrain data products;
- Identify key methods to refine the survey procedure and minimize errors through a series of controlled experiments;
- Develop and apply a refined methodology to undertake a further repeat survey of the Dart River in 2015 and assess the resulting data quality and compare with the earlier model;
- Identify the contributing factors leading to intrinsic and systematic errors within the SfM reconstructions;
- Develop a spatial representation of errors in the SfM derived models;
- Develop a means to mitigate systematic variations between the two surveys;
- Conduct analysis of change using the established Geomorphic Change Detection workflow (Brasington et al., 2003; Wheaton, 2010; Vericat et al., 2017).

1.7. Thesis Structure

The thesis is presented as a monograph comprising nine chapters, incorporating four key chapters detailing the empirical results. A brief outline of this structure is given below:

1.7.1. Chapter 1: Introduction

The rationale for the thesis is outlined through the identification of the requirement to derive geomorphological data sets over increasing spatial extents and at higher temporal frequencies. The practicality of deriving such datasets has been limited previously through the recourse to expensive equipment and methods of deployment that required extensive training and capital facilities. The advent of SfM provides a low cost alternative to derive topographic datasets at comparable point densities to that of traditional laser ranging. The use of SfM is proposed for capturing a 35 km reach of the Dart River, to assess the geomorphological evolution of the river in the immediate aftermath of large sediment input.

1.7.2. Chapter 2: Literature Review

This chapter outlines the fundamental theory underpinning SfM photogrammetry and draws comparisons to traditional softcopy photogrammetry. This necessitates an in-depth examination of the optical fundamentals of SfM, from the pinhole camera model, through to the estimation of the fundamental matrix. The typical pipeline used to generate 3D reconstructions is described, outlining the key steps involved in keypoint identification, object matching, bundle adjustment and densification steps taken to produce point clouds. The range of applications of SfM within geomorphology is explored, with attention paid to the current opportunities and challenges posed. Finally, the methods used to support the analysis of geomorphological change detection are described, considering both surface and point-based methods and appropriate error models.

1.7.3. Chapter 3: Site and Acquisition

A brief background to the geomorphological setting of the Dart River is presented. This identifies a simplified classification of the system into a series of spatially differentiated compartments, each with their own technical and logistical survey challenges. This classification of the reach is used throughout the thesis for simplicity of referencing.

The second half of the chapter outlines the data acquisition processes used in the research, focusing in particular on the large scale field methods. Finally, the chapter describes the photogrammetric pipeline employed within Agisoft Photoscan (hereafter APS), the consumer software system used throughout this research. The steps involved

in this software are linked to the fundamentals of SfM as detailed in Chapter 2 where possible, although it should be noted that the precise algorithms used are proprietary and only weakly documented.

1.7.4. Chapter 4: 2014 Survey

A SfM pipeline is developed and used to reconstruct a 35 km reach of the Dart River from the Slipstream debris fan to the Delta at Lake Wakatipu based on imagery acquired in May 2014. This survey was supported by an NERC Urgent Grant (NE/M005054/1). In order to meet the objectives of this grant, the survey was undertaken rapidly in order to provide an emergency assessment of local risks required by the local research partners, GNS Science and the Otago Regional Council. There were severe logistical challenges associated with this survey, in part forced by the urgency of data capture, and these are logistical constraints and their implications are discussed in detail. The resulting terrain products are initially validated through an analysis of Independent Control Points (ICP) errors, before comparing to an existing LiDAR derived DEM that was made available from the Otago Regional Council, extending over the southernmost region of the Dart delta. A further contemporaneous dataset for model evaluation was obtained using long-range TLS over the upstream debris fan. These dense datasets were used to develop a series of methods and metrics for model evaluation and to assess key aspects of the reconstruction processing steps in APS.

1.7.5. Chapter 5: Flume Survey

The evaluation of the rapidly acquired 2014 survey suggested that a number of improvements in the resulting terrain products could be made by refining aspects of the data collection procedures. In order to help design changes to the survey protocol, a series of controlled experiments were developed using a scaled-flume model. This research aimed specifically, to assess how modifications to the camera network geometry could help minimize systematic errors in the resulting block triangulation and potentially relax the reliance upon ground control. An optimal convergent imaging network is proposed which is supported by the use of a dual camera rig. The impact of image format and density is examined to find ways to mitigate the expected extra computational burden from this new multi-sensor approach.

1.7.6. Chapter 6: 2015 Survey

A repeat survey of the Dart River conducted in 2015 is reported. In this, the survey procedure was refined based on the insights derived from the experiments outlined in Chapter 5. Logistical challenges arising from the use of low-cost camera positioning tools and the role of convergent imagery in pose estimations are discussed. An evaluation of the terrain products derived is once again assessed through the analysis of ICP values and spatially continuous data derived from a low-altitude UAS survey of the Hillocks sub-reach. Significant improvements in the model quality metrics are identified.

1.7.7. Chapter 7: Geomorphological Change

The data requirements for reliable geomorphological change detection are discussed and the derivation of a spatially-explicit error model based on a minimum level of detection (minLOD) for different surface types (covers) is proposed. DEM differencing using the Geomorphic Change Detection (GCD) workflow highlights major confounding systematic errors. These systematic errors are found to originate principally in the 2014 dataset and a series of methods to empirically correct this model are examined. Results are presented that describe the observed geomorphological change along the lower 11 km reach of the Dart River between 2014 and 2015 for exposed bars.

1.7.8. Chapter 8: Discussions

The broader findings of this thesis are considered; specifically the results of using convergent imagery and prior pose estimations to minimize systematic errors. A small number of potential improvements are suggested based upon both surveying practice and processing, before a discussion on the importance of the choice of error metric when reporting survey accuracy.

1.7.9. Chapter 9: Conclusions

The findings of the thesis are compared to the main research objectives and the future potential of catchment-scale SfM studies is briefly covered, specifically, the need to begin using SfM data products for the purposes of geomorphological change detection.

2. Chapter 2: Literature Review

Abstract

The key research questions explored within this thesis focus around the construction of catchment-scale DEMs derived using SfM Photogrammetry, principally for the purpose of quantifying fluvial morphology and geomorphic change detection. The following chapters examine the emergence of SfM Photogrammetry, arising from developments in both computer vision and traditional soft-copy photogrammetry, before discussing the basic theory upon which the method rests. Following a presentation of the photogrammetric methods, the uptake and application of this new approach within the geomorphological literature is reviewed. This analysis examines the opportunities SfM presents and the challenges it poses, leading to areas for on-going and further research. Finally, the chapter discusses the quantification of geomorphological (topographic) change based on the comparison of sequential terrain datasets through time. This review examines approaches that use different data models, specifically raster DEMs and 3D point clouds, and explores the methods used to segment predicted changes likely to be due to actual geomorphological activity from those arising spuriously from model errors.

2.1. Structure from Motion Photogrammetry

2.1.1. **From manual measurements to automated software**

The use of analytical photogrammetry in geomorphology dates back to the mid-19th century, with documented applications relating to mapping Alpine environments in the 1890's (see Collier, 2002). With the improvements in aerial image capture, photogrammetry has long been recognised as an important method to generate topographic information (Chandler, 1999; Lane, 2000). Despite the potential power of photogrammetry, until the last five years, there has been a relatively low level of adoption within the wider geomorphological community. In part, this reflects the complexity of the optical theory, with Lane (2000) highlighting the need for an improved understanding of the relevant parameters, data quality and general training in photogrammetry. More recently, Baily et al. (2003) noted uptake was sensitively dependent on the benefits of increased automation, which improves accessibility and removes elements of subjectivity. The emergence of Bundler (http://www.cs.cornell.edu/~snavely/bundler/) and Microsoft Photosynth (https://photosynth.net) as free, semi-autonomous, web-based tools that implement the key initial steps in the SfM workflow has dramatically increased the popularity of photogrammetric methods in fields ranging from geomorphology, archaeology, engineering, architecture and city modelling (Fonstad et al., 2012; James and Robson, 2012; Rosnell and Honkavaara, 2012; Verhoeven et al., 2012). There now exists a plethora of subscription (commercial), freeware and open-source software including Agisoft Photoscan (APS), Autodesk 123D, Mic Mac, Pix4D, VisualSfM (see Stumpf et al., 2015; Westoby et al., 2012; Fonstad et al., Gomez-Guittierez et al., 2014a; 2014b; Kaab et al., 2014; Turner et al., 2012; Woodget et al., 2015; see further details in section 2.5).

Whilst promoting uptake, there are latent dangers in simplifying and automating SfM, especially where visually attractive results (photo-rendered point clouds) may detract attention from significant systematic errors that originate from poorly designed photogrammetric camera networks or less than robust ground control. Furthermore, many of the most popular software (e.g., Agisoft Photoscan or APS) now used for SfM are based on proprietary codes which, due to commercial sensitivity are poorly documented. The result is a relatively opaque processing chain, which hampers identification and understanding of the sources of error and sensitivity of parameters.

This is compounded by the standard treatment of errors in SfM models, which are largely reported spatially, as averaged residual values of independent control points (Lucieer et al., 2014a). Comparatively few studies make use of alternative spatially continuous data to facilitate spatially explicit comparisons, for example using independent laser ranging datasets (Westoby et al., 2012). There are encouraging signs, however, that recent studies (James et al., 2017) are beginning to develop error models based upon photogrammetric precision estimates.

To support a greater understanding of the processing workflow and potential sources of errors and uncertainties, this review first focuses on the fundamentals of SfM Photogrammetry.

2.1.2. **Theoretical basis of SfM**

Techniques to reconstruct 3D models from photographic imagery have undergone a revolution in the past decade, enabling a step-change in the quality of topographic data to support geomorphological research (James and Robson, 2012; Westoby et al., 2012; Lucieer et al., 2013; Solbo and Storvold, 2013). While such 3D reconstructions from photography have traditionally relied on standard photogrammetric principles (e.g., Lane 2000; Brasington et al., 2003), advances in the last ten years have stemmed largely from developments in computer vision. Building on the theoretical foundation of Marr and Poggio's (1979) seminal paper on stereo vision, image-based 3D reconstruction methods have been transformed by advances in computer power, feature tracking and bundle adjustment algorithms. This has enabled the semi-automatic generation of 3D point cloud data from highly redundant (overlapping) imagery. The state-of-the-art processing chain, or pipeline, is today most commonly known as Structure-from-Motion or SfM Photogrammetry, see figure 2.1. Unlike traditional photogrammetry, which relies on manual calibration of both the interior and exterior camera orientation, SfM software provides a means for self-calibration; the process by which internal camera parameters are estimated from a set of arbitrary un-calibrated images (Westoby et al., 2012).

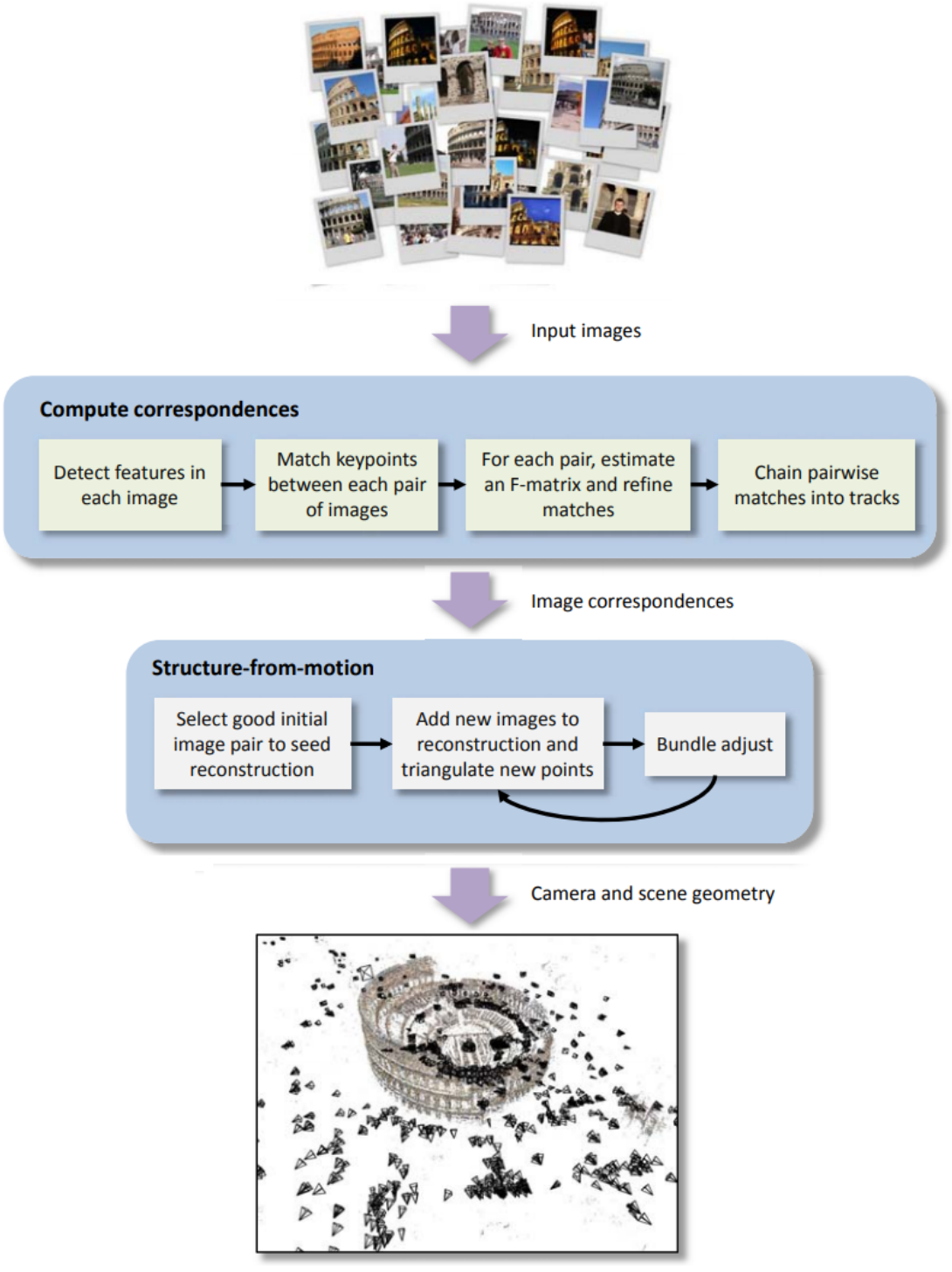

Figure 2.1 Visualisation of the SfM pipeline as a means to reconstruct scene from unordered online photosets. It's important to note the number of processing steps required, which results in the potential to use multiple combinations of algorithms (figure source, Snavely et al., 2008; p. 28).

The SfM pipeline seeks to solve a solution to both the internal and external camera orientation simultaneously, increasing the number of parameters to be solved, but

presenting the opportunity to incorporate compensation between the two. Such an approach is advantageous due to the lack of requirement of a calibration target – a potentially onerous task – and by virtue of the fact that camera parameters may vary during surveys due to vibrations or thermal and mechanical shocks (Civera et al., 2009). Self-calibration calculates the projective parameters: focal length, principal point and skew which is used to inform the initial conditions for bundle adjustment. This approach comprises a collection of processes built on the simple model of a pinhole camera to represent image projection and infer scene structure from multiple views or camera motion (Rastgar, 2013).

The SfM "problem" is commonly broken into two components; the measurement of 2D image displacement/velocity – known as correspondence or optical flow, and the extraction of 3D motion and structure using only 2D inputs (Xiang and Cheong, 2003). Given *n* points projected on *m* images, SfM identifies the projection matrix *P* and consistent structure *X* (Marcis, 2013). Briefly summarised, SfM involves the estimation of the Fundamental Matrix, **F** or Essential Matrix, **E**, based on a sufficiently large number of correspondences between images (Xiang and Cheong, 2003; Ni and Dellaert, 2012). This problem, however, remains poorly posed with no single, general purpose algorithm to emerge as yet. As such, SfM typically involves a sequence or pipeline of methods (Gemeiner, 2013; Rastgar, 2013). The implications of such a pipeline are twofold. First, it facilitates a high degree of user customisation, enabling careful choice over suitable algorithms for each stage of the process, but simultaneously, this very complexity frequently leads users to adopt a black-box approach with insufficient understanding applied to processes that might generate significant systematic and random errors.

2.2. Pinhole camera model

SfM originated in stereo vision from the pinhole camera projection model that describes imagery obtained by travelling about a point at a known velocity (Xiang and Cheong, 2003). By contrast, traditional photogrammetry is based upon the derivation of measurements from two separate cameras. The pinhole model can be seen as an approximation of the behaviour of real world cameras, without the additional complexity posed by more complex internal camera geometry and lens aberrations

(Robertson and Cipolla, 2009). The model delineates the relationship between a 3D real world point and its corresponding 2D point upon the image plane through three steps (Rastgar, 2013). The initial step is to transform a point X in the real world co-ordinate system to X_C in the camera co-ordinate system, through the similarity matrix:

$$X_C = \begin{bmatrix} \boldsymbol{R} & \boldsymbol{t} \\ 0 & 1 \end{bmatrix} X \qquad [2.1]$$

where $\boldsymbol{R}$ is the rotation representing camera orientation and $\boldsymbol{t}$ is the translation vector between camera and world co-ordinates. These are known as the extrinsic camera parameters, representing camera pose (Robertson and Cipolla, 2009; Rastgar, 2013). The next step requires the 3D to 2D transformation of the point X_c in the camera co-ordinate system to point x_c in the image plane co-ordinate system:

$$x_c = K[I \quad 0]\, X_c \qquad [2.2]$$

where I is the identity matrix and K represents the intrinsic camera properties, as represented by the upper triangular calibration matrix:

$$K = \begin{bmatrix} f_x & s & u_c \\ 0 & f_y & v_c \\ 0 & 0 & 1 \end{bmatrix} \qquad [2.3]$$

where f_x and f_y are the focal lengths in both the x and y co-ordinates respectively, s is the skew – which is assumed to be 0 as most modern cameras have perfectly perpendicular sensors (square pixels) – and (u_c, v_c) represents the optical centre in the pixel co-ordinate. The final transformation required is simply from the image plane co-ordinates to image pixel co-ordinates. All three steps can be combined conveniently into a single linear equation:

$$x \sim PX \qquad [2.4]$$

where $P \sim K\,[\boldsymbol{R} \quad \boldsymbol{t}]$ and is a 3 x 4 projection matrix. It is worth noting that real world cameras will differ from the pinhole projection via a displacement between points due to lens distortion; most frequently the non-linear effect of radial distortion (Robertson and Cipolla, 2009). Traditionally, photogrammetry typically uses metric cameras with well-known (a priori) camera models to minimize such additional complexity. Within many SfM routines, these unknowns (refinement of focal length, radial distortion,

principal point offset, tangential distortion, affinity and orthogonality) are, by comparison, incorporated into the bundle adjustment. This approach facilitates implicit self-calibration of the camera and the use of cheap, compact cameras for image acquisition, although is not without potential drawbacks (James and Robson, 2014).

2.3. Fundamental and essential matrix

The geometric theory of SfM infers scene structure and camera motion from correspondences between features in multiple images (Ni and Dellaert, 2012; Marcis, 2013). The inherent projective geometry between multiple views is known as the epipolar geometry and is encapsulated by the fundamental matrix (Hartley and Zisserman, 2004). The epipolar geometry of intersecting image planes acquired from two positions is shown in Figure 2.2, and is the foundation for corresponding point searches in stereo vision.

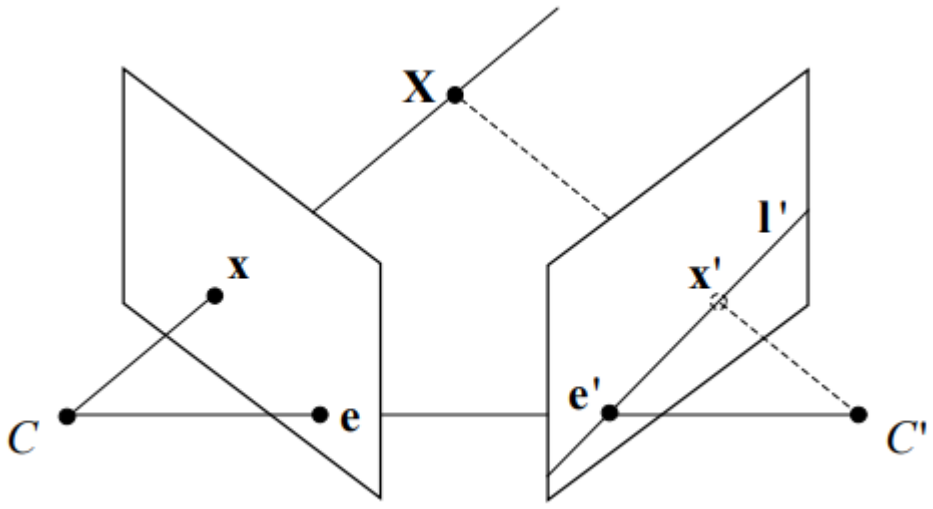

Figure 2.2 When observing point $\boldsymbol{X}$ from optical camera $\boldsymbol{C}'$, the projection of $\boldsymbol{x}'$ within the right image is restricted to the epiopar line $\boldsymbol{l}'$ The epipolar line $\boldsymbol{l}'$I is the image in the right camera of a ray through the optical centre $\boldsymbol{C}$ and image point $\boldsymbol{x}$ in the left camera. This is equivalent to intersecting the plane generated by the optical centres $\boldsymbol{C}$, $\boldsymbol{C}'$ and image point $\widetilde{\boldsymbol{x}}$ with the right image plane. The projection of the plane generated by intersecting each optical centre in each image represents the epipole $\boldsymbol{e}$. (Figure source, Robertson and Cipolla, 2009; p. 10).

The fundamental matrix is the algebraic representation of the epipolar geometry and serves to map points to lines across un-calibrated image pairs, $x \rightarrow l'$ (Rastgar, 2013), shown by:

$$l' = Fx \quad [2.5]$$

where F is the fundamental matrix and l' is the epipolar line in the second image which corresponds to point *x* in the first. Scene reconstruction is reliant on the accuracy of matching correspondences or features, identifiable in pairs or multiple images. The fundamental matrix provides an advantage in terms of algorithm construction which is evident in the reduction in computer processing required to search an entire image plane versus a single line (Hartley and Zisserman, 2004).

Typically, the fundamental matrix is calculated using a minimum of seven $x \Leftrightarrow x'$ correspondences; where every image point correspondence provides a single linear constraint over the nine coefficients of the fundamental matrix. Typically, more than seven correspondences will be available and so a least squares solution may be found for the matrix $\boldsymbol{Ab} = 0$, where $\boldsymbol{A}$ is a $n \times 9$ measurement matrix and $\boldsymbol{b}$ represents the 3 x 3 fundamental matrix elements f_{ij}. Through calibration, the fundamental matrix relates to the essential matrix through normalized image co-ordinates (Xiang and Cheong, 2003; Rastgar, 2013). Much like the fundamental matrix, the essential matrix maps points in one image to a line in another within the camera co-ordinate system:

$$\hat{x}^T E \hat{x}' = 0 \quad [2.6]$$

where $\hat{x}$ and $\hat{x}'$ are corresponding points within the camera system, E is the essential matrix and T is the matrix transpose (Hartley and Zisserman, 2004; Zhen et al., 2013). Essentially, given a number of corresponding points, it is possible to determine the fundamental matrices, from which the camera intrinsic properties, K can be calculated using self-calibration (as described in 2.2 above). The essential matrix may be decomposed to provide ***R*** and ***t*** – providing camera position and motion and as a result, the projection matrix (Robertson and Cipolla, 2009). Given known projection matrices, corresponding image points can be computed in 3D space via triangulation. Ideally, points will be located at the intersection of back-projected rays; however this is unrealistic due to the inherent noise in the process (Robertson and Cipolla, 2009; Rastgar, 2013). As such 3D points are chosen with a mind to minimize an error metric or cost function (Ni and Dellaert, 2012).

2.4. Software implementation

This study has been conducted with the professional edition of the package APS which breaks down the reconstruction process into a series of alignment, geo-referencing and densification steps with varying degrees of user interaction. However, as a black box toolset lacking documentation of the precise algorithms used, a number of inferences are necessary when interpreting the processing steps of APS in relation to the fundamentals of SfM described above. In order to avoid undue assumptions, this section shall outline the implementation of the generalized SfM pipeline, independent of software, while section 3.4 provides detail on how these steps relate specifically to the processes defined within APS.

The common components may be broken down as: (i) image and GCP acquisition; ii) keypoint identification; (iii) bundle adjustment; (iv) geo-rectification; and (v) pixel-based densification. These steps are summarized graphically in Figure 2.3.

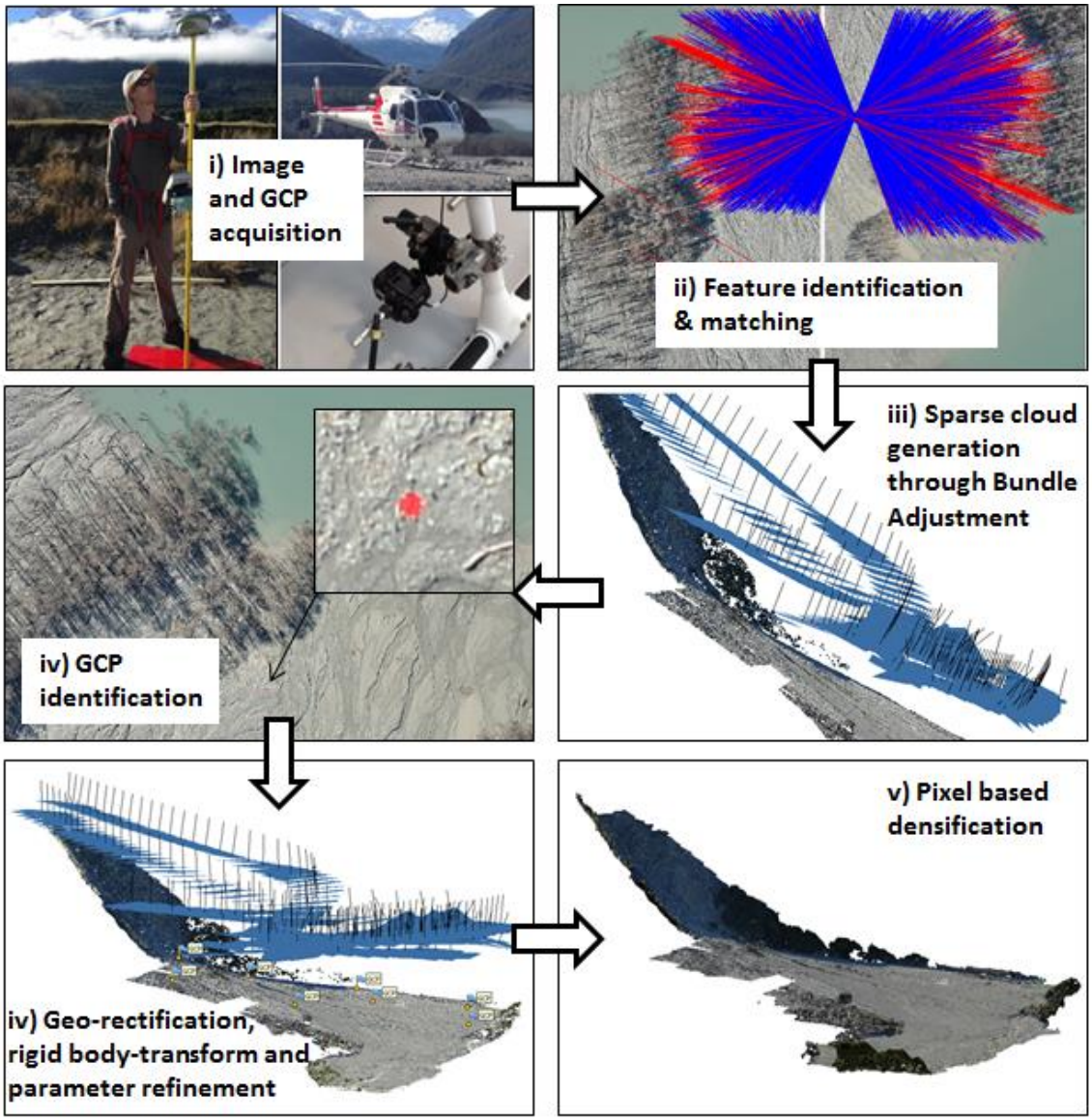

Figure 2.3 The SfM 'pipeline', where by i) image capture and the construction of a GCP network are the prerequisites. Once this information is obtained, ii) an automated software process derives and matches keypoints and iii) generates sparse clouds through the bundle adjustment. Subsequently, iv) manual identification of GCPs within the imagery allows a 7 parameter real world transformation to orientate the model, whilst further calibration refinement is possible through the improved accuracy introduced by these exterior constraints. This geo-referenced sparse cloud provides the camera positional data to v) reconstruct depth and generate the dense cloud through pixel matching.

2.4.1. **Keypoint identification**

Keypoint identification is the first process within SfM, and involves the detection of n distinct features within individual images that are then matched across pairs of images. The process of feature matching to support correspondence analysis is typically achieved using algorithms that aim to minimize the sum of squared errors between measured and predicted image locations. Numerous feature detection methods exist

such as SIFT (Scale Invariant Feature Transform), the Harris Corner Detector, SURF (Speeded Up Robust Features) and the Kanade-Lucas-Tomasi approach (Lowe, 2004; Govender, 2009; Rastgar, 2013). These algorithms typically examine neighbourhood variations in pixel intensities (brightness numbers) within a small window that may be shifted directionally and rotationally to determine the nature of the feature. The Harris corner detection algorithm provides a simple introduction into feature matching. In this context, a 'corner' is defined as a feature with large variations in pixel intensities around the neighbourhood in all directions. The algorithm examines the intensity values within a small window, which when shifted in direction leads to a large variation and is taken to indicate a corner point. This may also detect edges as moving the window yields no change in the edge direction (Harris and Stephens, 1988).

The Harris corner algorithm is rotation invariant; a corner remains a corner regardless of the rotation of the image; however corners identified in this form are sensitive to image resolution (scale), which may vary significantly within an irregular camera network. As such, the most commonly cited algorithm used in SfM application is the Scale Invariant Feature Transform or SIFT algorithm after Lowe (1999; 2004; see also Snavely, 2006; Govender, 2009). Lowe (2004) proposed SIFT as a five-stage process. The first step highlights regions of rapid intensity change, before detecting local extrema across scale and space to indicate the presence of a keypoint. The second step locates these keypoints and refines them via a Taylor series expansion, which increases accuracy of extrema, determining if they pass the requisite threshold, removing low contrast points and edges, see Figure 2.4.

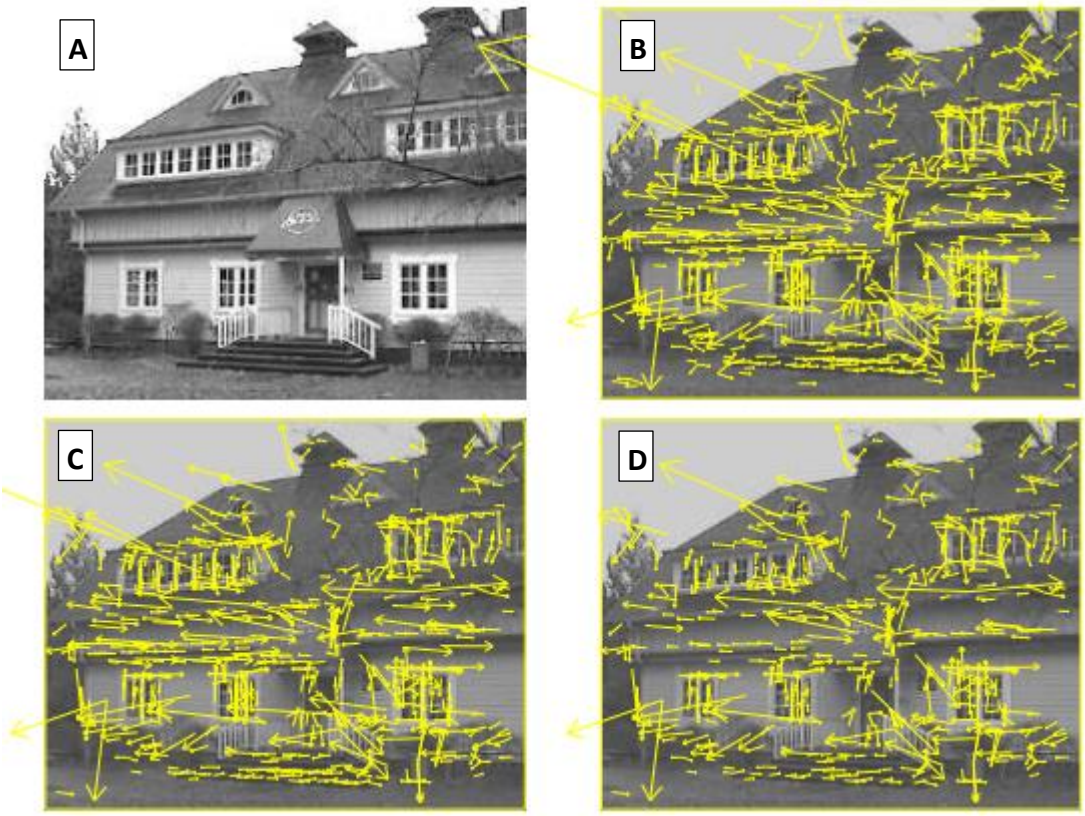

Figure 2.4 The stages of keypoint selection, with a) the original image; b) keypoint locations at maxima and minima of the difference-of-Gaussians function, displayed as vectors indicating scale, location and orientation; c) initial thresholding based upon contrast to reduce keypoint numbers and; d) final remaining keypoints after threshold applied to remove those poorly localized along an edge (figure source, Lowe, 2004; p.11).

The third step achieves invariance to image rotation through orientation assignment based on a threshold application to an orientation histogram; creating keypoints with same location and scale but different directions, which assists the robustness of the points identified. In the fourth step a keypoint descriptor is created in the form of a vector, which assigns a numerical description of the image gradient to all keypoints; a process that is robust to variations in brightness, contrast and viewing angle. The final step then matches keypoints between images.

The results of SIFT are commonly optimized using the RANSAC (RANndom SAmple Consensus) algorithm, which estimates parameters iteratively from a data set containing outliers – essentially refining the geometric variation using epipolar constraints. The refined matches are then used to estimate the fundamental matrix (Snavely et al., 2006; Civera et al., 2009). Refinements in feature matching algorithms have led to the creation of comparable alternatives, which in the case of the Speeded Up Robust Features or

SURF algorithm, trades computational efficiency and robustness, for the detriment of handling changes in viewpoint and lighting (Bay et al., 2008; Govender, 2009). The choice of algorithm used may, therefore, impact significantly upon results and this, in turn, may depend on the nature of the scene and features it incorporates in the imagery obtained. For example, within Photoscan the HEIGHT and EXACT algorithms have been specifically designed for surveying topography utilizing aerial photography (Verhoeven, 2011).

2.4.2. **Bundle adjustment**

The subsequent step in the SfM pipeline is the transformation of the previously identified keypoints to a sparse cloud. The bundle adjustment is the non-linear minimization of projection errors, which enables the construction of geometry of a 'sparse cloud' based upon feature matching (Snavely, 2006; Ni and Dellaert, 2012). This essentially finds the maximum likelihood estimate of a 3D point from a set of projection matrices (Rastgar, 2013). The name is a reference to 'bundles' of light rays hypothetically emitted from each object that converge on each camera centre and are then 'adjusted'. The adjustment refines the point cloud reconstruction to jointly produce optimal structure, camera pose and calibration estimates (Triggs et al., 2000). Accurate initialisation is required before to constrain this optimization process, as the adjustment acts as a sliding window which recursively provides camera pose and feature location estimations from the image dataset and is sensitive, therefore to local, sub-optimal minima in parametric space (Civera et al., 2009).

The problem is most commonly expressed as:

$$\min_{P_k, X_k^i} \sum_{ik} d(P_k X_i^k, x_i^k) \qquad [2.7]$$

where $d(x, y)$is the geometric distance. Various numerical methods can be used to minimize the cost function by bundle adjustment, including Newton-Raphson, the Gauss-Newton approximation and the Levenburg-Marquardt (LM) algorithm. As an example, the LM algorithm, known also as the damped least square method (Snavely et al., 2006; Robertson and Cipolla, 2009), and is a gradient-based iterative optimization method, which converges on minima initial estimates. Within each step the parameter vector is replaced by an updated estimate.

In the absence of metric cameras, the bundle adjustment is undertaken with no *a priori* knowledge of calibration data within APS. Previous studies have performed this through the use of calibration targets (Bertin et al., 2015), software such as Agisoft Lens (Clapuyt et al., 2016) through using test alignments from a subsection of images or through novel in-flight calibration via LiDAR control points (Gneeniss et al., 2015).

2.4.3. **Geo-rectification**

The resulting sparse cloud (in the absence of positional camera data) is an unscaled 3D point cloud in arbitrary space. For the purpose of analysing within a real world, universal coordinate system, an extensive network of accurately positioned Ground Control Points, hereafter GCPs, which are visible on the acquired image, are required. Subsequent identification of these marks within the image set and, therefore, the sparse cloud, enables the application of a standard seven parameter rigid body transform (e.g., Horn's method) comprising translation, rotation and scale (James and Robson, 2012; Rosnell and Honkavaara, 2012; Westoby et al., 2012) to transform the resulting point cloud to world space. A minimum of three targets across the survey extent are required; however, these do not need to be visible simultaneously in all images comprising the photogrammetric block unlike traditional photogrammetry (Smith et al., 2016). The distribution and number of targets is dependent upon the survey conditions, however, regular spacing throughout and incorporating peripheral regions has been found to reduce the number of targets needed (see later discussion in section 3.3) James and Robson, 2012; Javernick et al., 2014).

2.4.4. **Dense reconstruction**

The initial sparse cloud gives rise to the scene structure, via control point referencing and optimization strategies which determine the final intrinsic and extrinsic calibration parameters. However, most modern day applications of SfM use this structure to support the generation of a densified point cloud, with point spacing comparable to that of TLS and limited only by the ground-space distance of the raw image pixels. Densification of the point cloud can be achieved through a range of numerical methods, referred to as dense matching algorithms. Typically, these involve the computation of a number of image depth maps, either globally or locally, using either stereo pairs or multi-view stereo, to enable consistent pixel-wise matching of the scene. These depth maps typically contain errors due to the noisy and discrete nature of the input images

(Mellor et al., 1997), so refinement of accurate depth maps is crucial for the successful dense 3D reconstruction (Li et al., 2015).

Whilst there exists a litany of available dense matching algorithms, these may be broadly broken down in to a number of general concepts; Multi-view, Pairwise, Local and Global. These comprise some, or all of the following stages; (i) a matching cost computation; (ii) cost aggregation; and (iii) disparity computation and iv) disparity refinement (Scharstein and Szeliski, 2002). These steps act to create disparity maps for image pairs, which provide depth estimates based upon the degree of pixel change. The greater the degree of change in a pixel between views, the less depth, whilst pixels that undergo relatively little change are deemed to be further away. In relation to the above steps, i) a disparity space image, a 3D matrix (x,y, disparity) is generated, which may be from an image pair, or incorporate multiple views; ii) cost aggregation is applied to aggregate the pixel matching cost over a defined region to reduce noise and ambiguities; iii) stable pixels are identified using a winner takes all approach, across multiple disparity images through mutual consistency checking; iv) the produced disparity map can be prone to regions of discontinuity dependent on the texture and occlusions within the scene and so require refinement to smooth (Scharstein and Szeliski, 2002; Li et al., 2015).

Global approaches to the above typically skip the aggregation phase as the majority of work is undertaken during the initial phase, determining a disparity function that minimizes a global energy, aimed at improving surface regularity (Hirschmuller, 2008). In contrast, local approaches focus upon both the cost computation and aggregation phases, with the step of computing final disparities becoming trivial. This is because this method performs a local winner takes all approach at each pixel, however, it is only able to enforce the uniqueness of matches within the reference image. Whilst traditionally, local pairwise methods have been used to produce disparity estimates, recent advancements in computational resources have enabled their succession by the global approach (Galliani et al., 2015).

The difference in Multi-view and pairwise approaches is best addressed through the differences observed in the MVS (Furukawa and Ponce, 2010) and SGM (Hirschmuller, 2008) algorithms. In brief, the multi-view stereo algorithm performs stereo matching for all overlapping image pairs where, once outliers are accounted for, all corresponding

image coordinates of each object point are used for triangulation to calculate the final 3D coordinates. In contrast, the SGM algorithm is a multi-baseline approach to stereo matching between a base image and all match images where all stereo matching results are combined by selecting the median value of all disparities for each pixel (Bethmann and Luhmann, 2015).

2.4.4.1. ***The Multi-View Stereo approach to dense reconstruction***

The MVS approach builds on the triangulation of the sparse cloud to generate depth maps by the triangulation and refinement of additional rays. MVS filters noisy data whilst increasing the number of points reconstructed by a magnitude of two to three orders (James and Robson, 2012). Utilising the camera geometry solved as part of the scene structure, MVS constructs a full resolution point cloud in which nearly every image pixel has a corresponding 3D location. In this form, the reconstruction pipeline can been seen to reflect three key stages. Matching algorithms such as SIFT provide accurate correspondences; the SFM bundle adjustment uses the correspondences to estimate precise camera pose; and then MVS uses this structure to generate dense 3D models with accuracy nearly on par with laser scanners (Furukawa et al., 2010; Westoby et al., 2102). The algorithm generates large numbers of points by working over a grid of pixels, searching for the best match for each cell. The size of the grid can be controlled to determine the reconstruction density with regards to computational cost (James and Robson, 2012). MVS produces a depth map for each view which are then fused into a single cloud, eliminating noise by averaging over consistent depth and normal estimates (Galliani et al., 2015).

The development of MVS has coincided with a drive within computer vision to focus upon increasing the amount of images SfM programmes can manage simultaneously (Ni and Dellaert, 2012). Recently, this has been achieved through the use of the CMVS (Clustering views for Multi-View Stereo), which clusters images into groups dependent upon their point of origin (James and Robson, 2012; Westoby et al 2012). CMVS takes the output of SfM as an input, before decomposing the input to a set of image clusters. Both MVS and CMVS offer unprecedented geometric fidelity, however, they place a large demand on the calibration procedure (Furukawa and Ponce, 2010).

2.4.4.2. The Semi-Global Matching approach to dense reconstruction

The most relevant form of densification algorithm to this thesis is that of Pairwise Semi Global Matching (SGM) based on the work by Hirschmuller (2005; 2008). This primarily varies from traditional MVS by selecting a base image against which matching pairs are seeded and compared on a one-by-one basis, as opposed to all overlapping pairs (Bethmann and Luhmann, 2015). While documentation of densification strategy used in APS is limited, it has been proposed that a pairwise SGM approach is undertaken (Verhoeven et al., 2012; Dall'Asta and Roncella, 2014; Yan et al., 2016). Traditional methods of pairwise matching are fairly rudimentary in comparison to MVS, however the SGM approach minimizes the noise found in matching images based on just one epipolar line, which are prone to noise and discontinuities. This is achieved by extrapolating eight or 16 points of view (paths) across the image (Figure 2.5), where each path provides information on the cost to reach a pixel with a certain disparity. Each pixel and disparity is summed across all paths and subsequently the disparity d at pixel p is chosen by identifying the lowest cost (Hirschmuller et al, 2012).

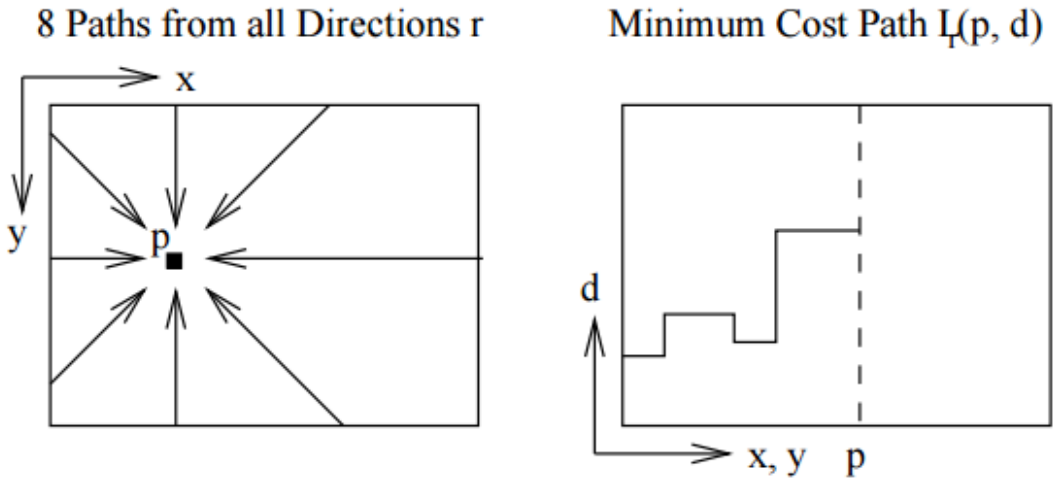

Figure 2.5 At each pixel $\boldsymbol{p}$ eight paths, $\boldsymbol{L_r}$ from all directions $\boldsymbol{r}$ meet. Along each path the minimum cost to reach all disparities $\boldsymbol{d}$ of a point is computed and visualised in the minimum cost path (figure source, Hirschmuller et al., 2012; p. 372).

Furthermore, an additional two tier penalty system is implemented, based upon identifying both surface changes and outliers. The system acts to penalise large local changes in pixel values, considered to be errors as opposed to natural change. The penalty system is applied during cost aggregation, in which disparities of at least 1 pixel are given only a minor penalty, while larger disparities are punished more aggressively. The summation of all points of view allows the identification of the minimal cost and

production of disparity maps for each stereo pair, which may then be fused to produce a final depth map (Figure 2.6).

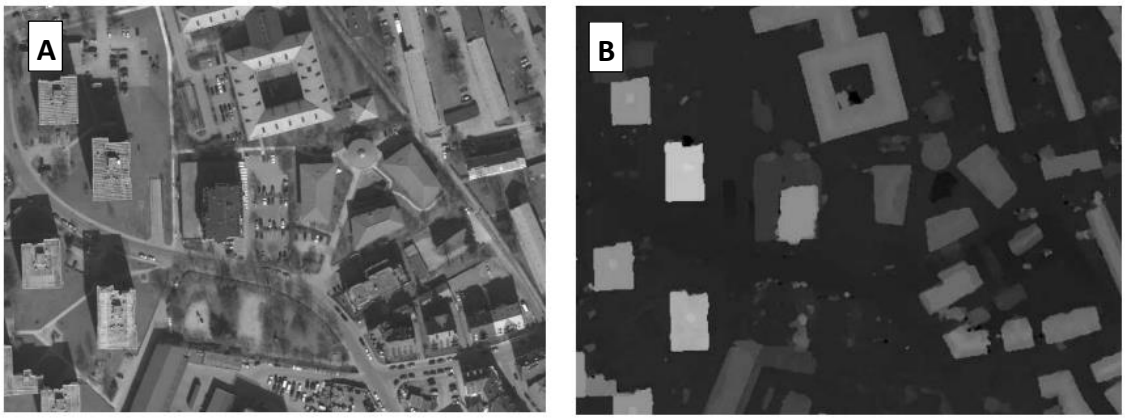

Figure 2.6 Example of a) orthophoto and b) depth map produced after implementation of the Pairwise SGM, whereby the colour graduation show's each pixels difference from the camera (figure source, Hirschmuller, 2005; p.7).

To produce a final depth map, disparity maps are orthographically projected using the known geometric camera model and fused to provide one consistent scene representation. This is achieved by projecting pixel values from all disparity maps on to a planar surface consisting of cells, whereby the median value for a pixel within a cell is selected. An advantage of such an approach is the ability to provide information from additional viewpoints that negates the need to interpolate between pairs in the initial disparity maps (Hirschmuller, 2008). The SGM method has been utilised within APS as the global minimization function is effectively reduced to an 8 (or 16) x 1D process and is, thus, significantly less computationally intensive than MVS-based densification. Furthermore, the structure of the algorithm is also suitable for parallelization (Hirschmuller, 2008; Bethmann and Luhmann, 2015) allowing solutions to make effective use of multiple CPU and GPU cores.

The importance of densification algorithms has been largely overlooked within the geomorphological community, where focus has largely concentrated on correcting the geometric errors associated with inaccurate camera calibration. As a result the repeatability of these differing densification algorithms has not been widely assessed. In the case of APS, the SGM algorithm functions independently of tiepoints within the sparse cloud, however is dependent upon the calibration and pose estimation to

determine the seeding order of stereo-pairs. As such, any variations in the sparse cloud and resulting calibration estimates are likely to proliferate into the densification process by altering the seeding order of stereo pairs and subsequent order of depth map construction and fusion. Furthermore, regions of poor surface texture are seen to be onerous, often requiring smoothing, as they initially result in streaky artefacts within the depth maps produced by the SGM (Dall'Asta and Roncella, 2014). It is important to assess the repeatability of these algorithms for areas of canopy and flat grassland, which may represent poor resolution, homogeneous surfaces within a high altitude high speed aerial survey (Haala and Rothermel, 2012).

2.5. Geomorphological Applications of SfM

2.5.1. **Trends within Geomorphological implementation**

While the theoretical and algorithmic developments underpinning SfM photogrammetry have been underway since the 1990s, the uptake of this approach beyond the computer vision community is a relatively recent phenomenon. It was not until the development of cloud-based tools, in particular Microsoft Photosynth, that the wider application of SfM became apparent in user communities beyond computer vision and close range photogrammetry. Within geomorphology, the adoption of this approach to surveying began in the current decade, with seminal papers introducing the approach emerging in 2012 (in particular, James and Robson, 2012; Westoby et al., 2012; Fonstad et al., 2013).

Since this time, SfM has been applied to monitoring landforms across a wide range of contexts and scales, covering fluvial, hill slope, coastal, glacial and volcanological settings. A summary of recent applications is presented in Table 2.1. Generally, these studies have demonstrated positive and encouraging results in terms of the quality of the data products produced. Effective reconstruction accuracies have been observed in a large number of studies, in particular where the distribution of ground control is dense and the spatial extent of survey remains relatively limited. The list presented in Table 2.1 is by no means exhaustive, but it is clear that to date, few studies (with notable exceptions including James and Robson 2012; 2014; James et al. 2017 and Clapuyt et al. 2016) have focused explicitly on identifying sources of reconstruction error and the

reproducibility of SfM data products. As such, there is an emerging dichotomy in the practically-oriented *application* of SfM Photogrammetry within the Earth surface science community and more analytical *assessment* of the methods being undertaken within the geomatics and photogrammetric literature.

Table 2.1 An overview of a recent number of varied geomorphological applications of SfM, the typical area covered and accuracy measure provided. The GCP and ICP accuracy values relate to 3D residuals unless stated otherwise. *Reference study from which other studies refer to SfM-MVS workflow.

Study	Geomorphological Application	Platform – Area- Software	Accuracy Recorded
Hugenholtz et al. (2013)	Aeolian Landforms	UAS capture of 1.95 km^2 are processed in Trimble Inpho.	ICP RMSE of 0.29 m
Snapir et al. (2014)	Agricultural	Hand-held capture of a 22 m^2 area processed in SfM Toolkit.	TLS based MAE of 0.004 m
Harwin and Lucieer (2012)	Coastal Environment	UAS capture of 0.0001 km^2 area processed with Bundler PMVS2.	ICP RMSE of 0.04 m
James et al. (2013)	Coastal Environment	Hand-held capture of 150 m^2 coastal cliff using SfM-MVS and 123-D Catch.	Majority within 0.002 m of TLS
Mancini et al. (2016)	Coastal Environment	UAS capture of 0.04 km^2 area processed in APS	TLS based RMSE of 0.22 m
Ruzic et al. (2014)	Coastal Environment	Hand-held capture of a 500 m stretch of cliff processed in Autodesk.	67% of ICP MAE within 0.075 m
Johnson et al. (2014)	Fault Zone topography	Balloon capture of 1 km^2 earthquake scarp processed with both SfM-MVS and APS.	90% within 0.41 m of LiDAR
Hudzietz and Saripalli (2011)	Flat Topography	UAS capture of 0.03 km^2 area processed in 3DM Analyst.	N/A
Brunier et al. (2016)	Fluvial Morphology	UAS capture of 150 m reach of river processed in APS.	ICP RMSE of 0.075 m (Z-only)
Dietrich (2016)	Fluvial Morphology	Helicopter capture of 32 km length of river processed in blocks within APS.	ICP MAE of 0.53 m
Dietrich (2017)	Fluvial Morphology	UAS capture of 250 m reach of river processed in	GCP MAE of ~0.04 m
Fonstad et al. (2013)	Fluvial Morphology	Helikite capturing a 0.036 km^2 area processed in Photosynth.	ICP MAE of 0.21 m
Jaud et al. (2016)	Fluvial Morphology	UAS capture of 0.07 km^2 area using both APS and MicMac.	ICP RMSE of 0.17 m (APS) and 0.078 m (MicMac)

Javernick et al. (2014)	Fluvial Morphology	Helicopter capture of a 1 km^2 area processed in APS.	ICP MAE of 0.34 m
Marteau et al. (2017)	Fluvial Morphology	UAS capture of 0.013 km^2 area processed in APS.	ICP MAE of up to 0.056 m
Smith et al. (2014)	Fluvial Morphology	Hand-held capture of a 600 m^2 area processed in APS.	dGPS MAE of up to 0.26 m
Tamminga et al. (2015)	Fluvial Morphology	UAS capture of 0.35 km^2 area processed in EnsoMOSAIC.	ICP RMSE of up to 0.169 m
Woodget et al. (2015)	Fluvial Morphology	UAS capture of 0.005 km^2 area processed in APS.	dGPS of 0.2 m (0.08 m with no vegetation)
Gomez-Gutierrez et al. (2014a)	Glacial	Handheld capture of a 0.004 km^2 area processed in 123-D Catch.	TLS based MAE of 0.21 m
Immerzeel et al. (2014)	Glacial	UAS capture of 1.75 km2 area processed with Bundler PMVS2.	Majority of GCP residuals under 0.2 m
Solbo and Storvold (2013)	Glacial	UAS capture of a 20 km^2 area processed in APS.	N/A
Tonkin et al. (2014)	Glacial	UAS capture of 0.2 km^2 area processed in APS.	Total Station RMSE of 0.517 m
Gimenez et al. (2009)	Gully Morphology	Kite capture of 0.04 km^2 area processed in LeicaGeosystem's LPS software.	GCP RMSE of 0.016 m
Kaiser et al. (2014)	Gully Morphology	Hand held capture of ~ 20 m^2 area processed in APS.	TLS based RMSE of up to 0.014 m
Lucieer et al. (2014a)	Landsliding	UAS capture of 0.008 km^2 are processed in APS.	ICP RMSE of 0.12 m
Niethammer et al. (2012)	Landsliding	UAS capture of 0.2 km^2 area using VMS close range photogrammetry software.	TLS based RMSE of 0.31 m
Kaab et al. (2014)	Periglacial	Hand-held capture from a 3 m ladder of ~25 m^2 processed in MicMac Apero	N/A
Lucieer et al. (2014b)	Periglacial	UAS capture of 0.02 km^2 area processed in APS.	ICP RMSE of 0.042 m
Turner et al. (2012)	Periglacial	UAS capture of 0.006 km^2 area processed in Bundler.	GCP based MAE of 0.103 m
Clapuyt et al. (2016)	Sloping Pasture/Arable Land	UAS capture of 0.01 km^2 area processed using Agisoft Lens and Visual SfM.	ICP MAE of 0.24 m
d'Oleire-Oltmanns et al. (2012)	Soil Erosion	UAS capture of 0.07 km^2 and 2 km^2 area processed in Leica Photogrammetry suite.	GCP residuals of 0.03 m and up to 1.7 m

Smith and Vericat (2015)	Sub-humid Badlands	Hand-held and UAS capture of up to 30 m^2 processed using APS	TLS based MAE of up to 0.18 m
Westoby et al. (2012)	Varied: Coastal Cliff, Breached Moraine Dam and Bed Rock Ridge	Hand-held capture of range of areas processed with Bundler.	61% of Vegetation free surface within 0.1 m of TLS
James et al. (2017)	Varied: Erosional Gulley and Landslide	Piloted gyrocopter capture of 0.08 km^2 gulley network and 0.2 km^2 landslide, processed using APS.	ICP RMSE of 0.047 m and 0.1 m
James and Robson (2012)	Varied: Summit Crater and Coastal Cliff	Micro-light aircraft capture of 1 km^2 area processed with SfM-MVS*.	Photogrammetric RMSE of 1 m, TLS of 0.1 m
Mathews and Jensen (2013)	Vineyard Canopy	UAS capture of 0.02 km^2 area processed in APS.	GCP RMSE of 0.166 m
Bretar et al. (2013)	Volcanic terrain	Hand-held capture of a 12 m^2 area processed in MicMac Apero.	N/A
James and Robson (2014)	Volcanic terrain	Hand-held capture of ~400 m^2 area processed using APS.	ICP RMSE of 0.11 m

In this tabulation of recent published studies, the area of application was chosen as a descriptor, primarily as the relationship between height and precision (a ratio of 1:1000) has already been explored in numerous previous papers (James and Robson, 2012; Woodget et al., 2015; Smith and Vericat, 2015). By contrast, distinguishing studies in terms of their areal extent incorporates the additional constraints of flight speed alongside height. Surveys covering larger areas typically involve a trade-off between both these aspects of data capture, and each have key photogrammetric implications, relating principally to pixel resolution and sharpness.

Table 2.1 illustrates a number of trends that dominate in existing geomorphological research to date. First, the most commonly used software is APS, reflecting its ease of use (no requirements for complex scripting), effective graphical interface, low cost, and community support. The dominant survey platform is UASs, which explains in turn the typically small extents covered (of the order 1 x 10^3 m^2) and high relative reported accuracies of between 0.1-0.3 m. By contrast Solbo and Storvold (2013) and Dietrich (2016) have undertaken much larger extents, exceeding 10 km^2, although independent accuracy assessments are not provided. Finally there is a lack of consistency in the reporting of error metrics (see section 3.7 for further description), with ICP and GCP

RMSE's (Root Mean Squared Error) and MAE's (Mean Absolute Error) most commonly reported, alongside the percentage of observations falling within a given vertical or 3D distance from control observations. Frequently, there is also lack of information in terms of the distribution of ICP locations, in particular their (potentially biased) proximity to GCPs and spread throughout the distal margins of the photogrammetric block. This is critical, for while aspatial mean error values provide an indication of overall model quality, aggregation of 'model performance' into such single values may offer little insight into key systematic errors, such as deformation at peripheral edges of a reconstruction which may reflect errors in the internal camera calibration. Spatially explicit measures, such as the percentage of the reconstruction lying within a given vertical confidence interval, may also fail to reveal the presence of extreme local residuals. As discussed below in section 3.7., what emerges from this practice is the need for a multi-metric approach to model assessment that identifies the influence of random and systematic errors as well as blunders in the resulting data products.

2.5.2. **Survey Platforms**

Whilst SfM shares a legacy with conventional aerial photogrammetry, the cameras and photogrammetric networks used differ significantly. For example, standard aerial photogrammetry typically relies on imagery obtained with large format metric cameras, acquiring data in nadir-orientated, parallel flight lines with typical endlap and sidelap of 55-65% and 20-40% respectively. By contrast, applications of SfM are often based on irregular camera networks, incorporating both vertical and oblique photography with very high overlap (60-90%) and generally acquired with low-cost, compact cameras. Indeed, a driving impetus behind the seminal work of Snavely et al. (2006) was to develop a SfM pipeline capable of using crowd-sourced images collected online and thus incorporating multiple camera models and a wide range of object-to-camera distances and viewing angles.

In principle, therefore, SfM is suited to photography acquired from a wide range of platforms, both aerial and terrestrial. In practice, however, the type of platforms used must be tailored to the specific nature of the study site and imaging target, and there emerges an inevitable compromise between spatial coverage and pixel resolution (Smith et al., 2016). For example, terrestrial-based solutions such as pole photogrammetry are ideally suited to small areas extending over 10^1-10^3 m^2, at which

scale, sub-centimetre pixel resolutions are achievable (James and Robson, 2012; Westoby et al., 2012). By contrast, the majority of studies are largely based on aerial image capture, obtained at lower photo-scale but over potentially much wider areas (10^3-10^6 m^2).

Arguably, a principal driver underpinning the surge of interest in SfM has been the wide range of flexible aerial platforms that are now available to support low-altitude photography with lightweight navigation and imaging systems. These include the use of kites and helium blimps (Verhoeven, 2009; Fonstad et al., 2013), but more recently unmanned aircraft systems (UAS); more commonly referred to as drones, or unmanned aerial vehicle systems (UAVs) or remotely piloted aircraft systems (RPAS; Williams et al., 2017). These flexible platforms are operated under varying degrees of autonomy, from full manual piloting, to pre-programmed flight paths guided by GNSS and terrain data. From a geomorphological perspective, these new platforms have radically reshaped the approach to data acquisition, placing the timing and design surveys in the hands of geomorphologists and creating the opportunity to dramatically enhance survey frequency in line with scientific objectives (Lucieer et al., 2014a). As the key technologies continue to develop apace (increases in payload and flying time coupled with miniaturization of imaging sensors, improvements in lens and spectral sensitivity) the scope of these platforms is set to expand dramatically, albeit checked somewhat by growing concerns over safety (see Cunliffe et al., 2017).

Commonly employed civil and commercial UAS may be classified into two broad design categories: multi-copters and fixed wing aircraft. Multi-copters are the most popular design, due principally to their low cost and ease of deployment that necessitates relatively little training. They remain, however, highly sensitive to wind conditions which may dramatically alter attitude (roll, yaw and pitch) and the smaller, cheaper models are limited by short battery life and, therefore, limited spatial extent (Verhoeven, 2009; Rosnell and Honkavaara, 2012; Smith et al., 2016). Fixed wing UASs are typically more expensive and require a higher degree of training to fly, although in turn, they offer the potential to cover larger areas, with survey paths of up to 130 km (Solbo and Storvold, 2013). Both categories are subject to air traffic legislation (Cunliffe et al., 2017) which dictates maximum flight heights and line of sight rules which

ultimately compromise their ability to collect data over large spatial extents in populated areas.

To acquire photography over larger areas, a small number of studies have employed full-scale piloted helicopters to acquire dense photography for SfM. These platforms combine aspects of the high manoeuvrability, low-speed and low-flying heights of UAS with comparatively unlimited flying times and the option to acquire imagery at heights in between UAS (20-120 m) and typical fixed wing aerial surveys (2000-5000 m) (Javernick et al., 2014; Dietrich, 2016). While such applications come with increased costs of operation, the relaxed conditions for camera deployment including hand-held or skid-mounted rigs, compared with standard survey aircraft, enable the use of a wide range of cheaper commercial helicopters reducing flying costs.

2.5.3. **Survey Capture**

Geomorphological applications of SfM observe a number of 'rules' when capturing imagery and ground control, comprehensively reviewed by Smith et al. (2016), which briefly may be summarized as relating to image quality, density, angle and GCP distribution.

2.5.3.1. Sensor and Image Quality

Image quality is closely linked to the imagery texture, which is an attribute of not only the scene, but the effective scale, resolution and pixel fidelity as determined by the sensor (Micheletti et al., 2015). A range of sensors may be used from video stills, compact cameras, fish eye lenses, through to high resolution, full format digital SLR cameras (Westoby et al., 2012). Micheletti et al. (2015) compared the use of a Nikon D7000 against an IPhone 4 camera processing the resulting imagery with PhotoModeller and AutoDesk 123D to assess the impact of sensor quality. While median DEM errors (defined against a reference TLS dataset) were found to increase for the lower quality sensor, the results obtained with the smartphone camera were still capable of generating fit-for-purpose data products. The impact of resolution and fidelity increases in importance as camera-to-object distance increases and theoretically at least, offset the use of higher flying heights to cover wider areas (Westoby et al., 2012; Smith et al., 2016). In practice, however, the choice of sensor and image resolution is constrained by the computational resources available and the processing engine of the SfM pipeline

(Westoby et al., 2012). Increases in camera resolution scale as n^2, so that enhancements in image resolution can place dramatic burdens on processing time. There is an important trade-off to be found, therefore, between resolution, image extent and the geomorphological requirements of the data. In a recent study of coastal morphology, for example, Harwin and Lucieer (2012) found that using compressed (reduced resolution) imagery that generated point clouds with only 1/5 of the maximum achievable density, were nonetheless sufficient for their application and only required 1/11 of the comparable processing time.

The use of non-metric cameras necessitates some practical solutions to obtaining reliable imagery. For example, to avoid introducing multiple camera models into a single photogrammetric block, it is standard practice to fix the camera focus at infinity, (James and Robson, 2012; Lucieer et al., 2014a) and this may include a physical-strapping (tape) to mitigate against unintended movement on unstable platforms by wind and changes in speed which may destabilise the calibration (Rosnell and Honkavara, 2012). Further constraints are introduced when capturing at high speeds as shorter exposure times are required to avoid motion blur (Rosnell and Honkavara, 2012), whilst well-exposed imagery is a primary requirement to maximize information in the imagery for keypoint identification and correlation. Given the wide range of viewing angles commonly used in an SfM camera network, improved results are obtained during bright overcast days (as opposed to direct sunlight), which provide diffuse lighting conditions (Gimenez et al., 2009).

Reflective surfaces, homogeneous surfaces, and areas of shadowing act to further complicate keypoint identification, resulting in often spurious matches and increased noise within the sparse cloud generated (Verhoeven, 2009; Furukawa and Ponce, 2010). As such, densely vegetated surfaces and wetted channels, and similar mobile targets, often generate numerous erroneous matches, which impact upon the subsequent geometry and resultant camera calibration parameters (Niethammer et al., 2012; Westoby et al., 2012; Fonstad et al., 2013). Despite taking such listed precautions, distortions may persist from motion blur and de-focus, with a standard approach across studies to remove or mask these images from the analysis (Westoby et al., 2012).

*2.5.3.2. **Image Network***

2.5.3.2.1. Image Density and Overlap

A key requirement of SfM over traditional stereo-photogrammetry is the need for highly redundant camera networks, providing typically >60-80% image overlap and extending beyond the area of interest to ensure complete coverage (Furukawa and Ponce, 2010). In published accounts, the density of camera networks is typically reported as percentages of sidelap and endlap, or given the irregular nature of acquisition, just a directionless overlap percentage. Examples from Table 2.1 show a variation in the degree of density, varying from Hudzietz and Saripalli (2011) who used 60% endlap and 50% sidelap; Hugenholtz et al. (2013) using 65% overlap; Bretar et al. (2013) using 70% overlap; Woodget et al. (2015) using 80%; and Lucieer et al. (2014a) with densities exceeding 80-90% overlap. Of particular relevance to the research objectives here, is the recent helicopter survey reported by Dietrich (2016). This survey of a single thread channel, the Middle Fork John Day, Oregon, extended over a comparable 32 km reach and used imagery with c. 60% endlap. Unlike the surveys of the Dart River conducted in this research, this system is a narrow, confined channel requiring only one flight line. The review of practice in Table 2.1 reveals a general pattern of between 60% and 80% side and endlap, although it is notable that higher densities are often promoted. For example, Rosnell and Honkavara (2012) advocate the imagery with up to 90% overlap in contrast to a minimum coverage of 60% as proposed by James and Robson (2012) and Fonstad et al. (2013).

Image density may also be expressed as a baseline (distance between cameras) to ground distance ratio, *b:gd*. In this context, Marcis (2013) found that the ideal *b/gd* ratio when using APS occurred between 1:5 and 1:2, although tended to vary across surface textures. At a higher ratio than 1:2 surface noise is introduced as parallel imaging angles introduce increased error within the z dimension (Kyto et al., 2011), whilst below the ratio of 1:5 there is insufficient overlap resulting in holes. This loss of detail can be seen in Gimenez et al. (2009) and Westoby et al. (2012) whom attributed it to large camera to feature baselines. Further considerations are the practical and computational constraints on increasing image density as discussed above (Smith et al., 2016). Large photogrammetric blocks (exceeding 1000 images) are associated with major

computation requirements (see Harwin and Lucieer, 2012) and a common strategy employed to manage this is the deconstruction of the scene into multiple overlapping blocks (e.g., Dietrich). While this approach enables the reconstruction of large areas not manageable within in a single computational block, it does pose further problems associated with merging these sub-regions, particularly given the frequently observed distortion at the margins of individual blocks.

2.5.3.2.2. Imaging Angle

Reconstructions based on nadir-orientated (vertical) camera networks have previously been found to be prone to systematic distortions that take the form of a dome or bowl characterised by the distribution of constraining ground control (James and Robson, 2012). This form of error arises principally as a result of inaccurate depth modelling (Wackrow and Chandler, 2011) due to inaccuracies in the internal camera calibration, which most commonly manifest in errors in the estimation of the radial distortion parameters of the camera model used, typically a variant of the Brown-Conrady (Brown, 1966) seven parameter model (James and Robson, 2014). The presence of such errors reflects the difficulty of simultaneously solving the exterior and interior camera orientation in a single block bundle adjustment, which presents the opportunity for compensation between parameters of the total solution in the presence of multiple sub-optima in the cost-function parameter space. Such erroneous lens calibration may be partially mitigated through the use of calibration targets (Bertin et al., 2015) and test alignments from a subsection of images and novel in-flight calibration via LiDAR control points (Gneeniss et al., 2015). However, unconstrained self-calibration is advantageous, relaxing the need for such targets, and more fundamentally as vibration, as well as thermal and mechanical shocks, imply significant uncertainty in the internal camera geometry during its use (Civera et al., 2009).

An alternative strategy to reduce the likelihood of inaccurate camera modelling is through the application of convergent camera networks (Wackrow and Chandler, 2008; Wackrow and Chandler, 2011; Marcis, 2013; James and Robson, 2014; Woodget et al., 2015; Dietrich, 2016). This is not a new idea and Karara and Abdel-Aziz (1971) initially proposed convergent camera angles to minimize projection errors. This approach was later tested by Wackrow and Chandler (2008) using photogrammetric simulation. In this analysis, they found that for a basic stereo-model, two cameras at 15^{o}, creating a relative

imaging geometry of 30°, provided the most reliable recovery of the accurate camera parameters. Their simulations identified differing degrees of radial correction $\Delta r_{a'}$ and $\Delta r'_{a'}$ required by camera *a′* observing a point A from two offset nadir perspectives (see Figure 2.7 below). By contrast the difference in radial correction required in a 30° convergent case is much more limited. The divergence in correction applied in the normal case results in the estimation of the point A (found by intersection between image pairs) to lie above the X-Y object plane, giving rise to a classical doming error. Conversely, the convergent scenario results in accurate placement of A on the X-Y object plane, although the correction may still nonetheless result in erroneous lateral displacement of the point. Such effects can, however be mitigated by re-estimation of the exterior orientation.

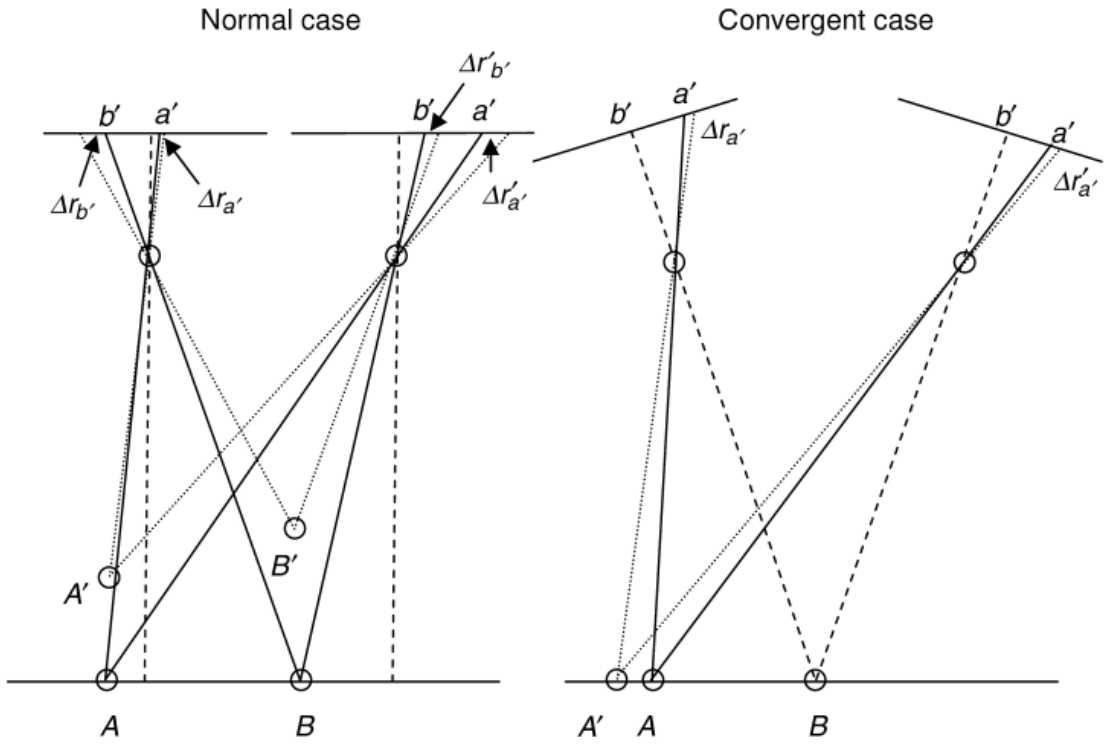

Figure 2.7 In the nadir orientation a), the radial distortion correction applied for object point A in the image space is significantly different across images ($\Delta\mathbf{r}_{\mathbf{a}'} \neq \Delta\mathbf{r}'_{\mathbf{a}'}$). In the instance where an inaccurate lens model is used, the intersection of rays from the left and right image occurs above the x-y object plane at$\mathbf{A}'$. In contrast, in b) the convergent network, the applied radial distortion correction is similar for both images ($\Delta\mathbf{r}_{\mathbf{a}'} = \Delta\mathbf{r}'_{\mathbf{a}'}$) resulting in the intersection of rays on the x-y object plane at $\mathbf{A}'$ however introduces a lateral displacement (figure source, Wackrow and Chandler, 2011; p. 22).

Wackrow and Chandler (2011) assessed the use of convergent imaging geometry empirically. Using stereo-photography acquired from a tripod mounted with pre-calibrated cameras, capturing a 4 m wide river channel with an object-to-feature distance of 8 m, they again found cameras angled at 30^{o} provided optimal viewing geometry. Replicating this approach for an aerial SfM survey is not straightforward. However, within a similar vein and using SfM (specifically APS), Marcis (2013) examined the impact of multi-angle sensors observing a planar surface. The error value was reported as a Standard Deviation against a reference dataset and as such incorporates both systematic surface deviations and noise. The optimum range for results was observed to occur for imaging angles between a nadir orientated camera and 45^{o} aft-nadir, with specific error minima at 25^{o}, similar to the 30^{o} observed by Wackrow and Chandler (2011).

2.5.3.3. ***GCP Distribution***

There is a very wide range in the practice of placing ground control in reported studies using SfM. The density of control used varies significantly, for example, Harwin and Lucieer (2012) used 90 GCPs over a 100 x 20 m stretch of coast; whereas Turner et al. (2012) used 20 in a similar area but covering an Antarctic moss bed. In contrast, Immerzeel et al. (2014) used 19 GCPs in their survey of 1.75 km^2 of a Himalayan glacier, while d'Oleire-Oltmanns et al. (2012) used 80 GCPs over a 2 km^2 landslide and Dietrich (2016) used only 68 points in a 32 km length of a single thread river system. This underlying variation in survey practice, in part reflects fundamental differences between SfM and traditional stereo-photogrammetry. In the latter, the number of GCPs required in a photogrammetric block can be minimized by manually measuring a larger number of quality tiepoints, and the use of a well-constrained metric camera. By contrast, in SfM the scene structure is solved initially using inner-constraints only (the keypoint correspondences). Ground control points are then used as exterior constraints to transform the arbitrary interior, to a world coordinate system. In principle, this could be accomplished by a rigid body transform requiring only 3-4 (with redundancy) constraints. However, in practice, GCPs are used to re-estimate the initial scene structure, either optimizing the interior camera parameters and/or iteratively updating the bundle adjustment. The density and distribution of control is critically important

where significant systematic errors, such as those associated with the lens distortion described above, may arise in the initial scene and GCPs are used to inform adjustments to the solution.

The lack of consistency in GCP density and placement, should be recognized as largely a result of practical concerns. The process is time consuming (d'Oleire-Oltmanns et al., 2012) and requires additional survey experience to position targets accurately and precisely using GNSS or total station measurements. Not surprisingly, therefore, considerable discussion has been devoted to identifying optimal densities and placement strategies to minimize this effort. Turner et al. (2012), for example, found that they were able to halve the number of GCPs used in their survey with minimal detriment to quality in the resulting reconstruction. Such reductions in density must, however, take account of the distribution of GCPs across the area of the model. In this context, Johnson et al. (2014) in reconstructions of desert topography using kite, blimp and glider photography, found that DEM quality was enhanced when GCPs were located in regions of maximal image overlap. Elsewhere, James and Robson (2012) cautioned that GCP placement should fully-encompass the area of interest in order to provide adequate mitigation of systematic errors within the study area.

Finally, and commonly reported, it is important to recognize that errors in the accurate placement and identification of GCPs will ultimately propagate into the reconstructed block. Such errors include easy-to-spot blunders associated with the misidentification of the targets, but also more subtle uncertainties such as the use of inconsistent datums, movement of GCP targets, and the incorporation of survey traverse errors. Such errors were found by both Clapuyt et al. (2016) and Dietrich (2016) to propagate linearly into the final reconstruction.

2.5.4. **Conclusions and challenges**

Structure-from-Motion provides a low cost, reproducible method for modelling topographic data across various geomorphological settings and has generated favourable results. Typical workflows for SfM have evolved principally from UAS based studies of small areas, characterized by dense ground control, and low flying heights. The field of SfM is constantly evolving, however, and perhaps the most exciting opportunities lie in the potential to map large areas at accuracies, and over temporal

resolutions, that have been previously unobtainable (financially and logistically). Whilst only a limited number of studies have attempted such large extents (Solbo and Storvold, 2013; Dietrich, 2016) there exists considerable scope to build upon these through improved workflows to minimize errors and to assess the reproducibility of these approaches. Dietrich (2016), for example, used nadir photography comprising >2000 images, which were reconstructed in a number of individual blocks to cover a 32 km river reach. There is therefore the opportunity for improvements to be made through increasing imaging numbers, capturing with additional oblique imagery and reconstructing the area in one photogrammetric block to minimize the need to merge individual reconstructions.

Following the generation of such broad-scale models, the logical progression is to assess the accuracy and reproducibility of these large scale datasets for key geomorphological applications such as topographic or geomorphological change detection. Recent work by Clapuyt et al. (2016) has attempted this by formally examining the internal and external uncertainties associated with the SfM pipeline. However, this study was again based on UAS photography over a small area of relatively simple terrain. The ability to derive geomorphological signals over regional or indeed catchment scale survey extent is currently unprecedented.

Such applications of SfM at large scales are clearly hostage to both local random and global systematic errors that affect DEM quality and, by inference, the reliability of topographic change estimates based on DEM differencing. These errors are likely to incorporate localised noise associated with key-point quality, which is sensitive to flying height and speed, alongside those generated by complex features such as vegetation and wetted channels. Systematic errors, in turn, are likely to result from sub-optimal camera calibration and weak ground control distribution; the latter being characteristic when working in remote, hostile environments over large areas. Whilst Wackrow and Chandler (2011) and James and Robson (2012) suggested that systematic errors typically manifest as a dome (or bowl), it is important to recognize that these studies have been based on regular or well-defined camera networks with simplified geometries. How such effects would be evidenced using alternative camera networks, such as those associated with corridor mapping, is less clear.

Finally, for SfM modelling to be established and trusted within the wider community, there is a pressing need to develop a transparent approach to model assessment, that combines not only simplified metrics of model fit, but also spatial measures that capture the effects of outliers, overlap effects, and systematic errors in the photogrammetric block. Importantly, there is a need to extend this error analysis to consider the fitness-for-purpose of the datasets generated. This must identify the robustness of predicted or derived quantities, such as volumes of geomorphological change, or the sensitivity of geophysical models to the inherent DEM quality.

2.6. Topographic Change

2.6.1. **Introduction**

The quantification of bed load transport in gravel bed braided rivers is notoriously difficult, given the high variability of transport rates in space and time, and the logistical difficulties of sampling over wide, complex braid plains at high flow. Over the last two decades, the potential to estimate transport inversely by quantifying the sediment budget and back-calculating the time-space integrated flux has received significant attention (Brasington et al., 2003; James et al., 2012; Vericat et al., 2017). This process has been dramatically facilitated by quantifying the morphological change between repeat topographic surveys by differencing directly either continuous surface models, such as raster DEMs or TINs or through comparison of raw point clouds (Lague et al., 2013). While the goal of such analysis is often cited as the inverse calculation of sediment flux, the basic process of DEM differencing provides more tangible information on changes in channel capacity (flood conveyance), gravel storage and availability and, importantly, provides spatially explicit mapping of how river morphology is adjusting and where.

The derivation of geomorphological change from terrain data must, like any form of comparison, account for sampling, measurement and interpolation errors that are inherent to the input models (Fisher, 1998) and propagate inevitably into measures of difference. Quantification of errors provides a vital first step in defining a framework to identify changes between two models that are statistically robust, that is, are unlikely to have occurred due to uncertainties in the data models. The ability to detect significant geomorphological change hinges upon the ratio of the observed signal to the combination errors into the input datasets (Brasington et al., 2003; James et al., 2012; Leary et al., 2012; Wheaton et al., 2013). A simple and common procedure for handling the measurement of change is to apply a minimum level of detection threshold (a minLoD) (Brasington et al., 2003; Lane et al., 2003), above which all observed differences are treated as statistically robust and below which are discounted as 'noise'.

The measurement of geomorphological change is a four dimensional process – x, y, z co-ordinates and time – with cartometric studies of volumetric change dating back to the mid 1970's (Brasington et al., 2000; James et al., 2012). The past decade has seen an

explosion in high resolution data capture techniques but comparatively little advance in the tools to analyse change. An exception to this, is the Geomorphic Change Detection methodology for raster DEMs that has been incorporated into the GCD software (http://gcd.joewheaton.org/) and is provided as a plug-in for ArcGIS (Wheaton et al., 2010). The analysis of change through point cloud datasets is still in its comparative infancy but holds promise as, by working with the raw observations, there is less potential for the incorporation of interpolation error (Barnhart and Crosby, 2013; Brasington et al., 2012; Earlie et al., 2013; Lague et al., 2013). Nonetheless, change detection from dense capture techniques remains difficult due to the complexity and density of point clouds which require extensive data cleaning and decimation and then the development of computational geometry rules to match point clouds across temporal epochs (Brasington et al., 2012; Barnhart and Crosby, 2013).

2.6.2. The DEMs-of-Difference approach

The geomatics revolution has enabled the acquisition of topographic data and modelling of DEMs at timescales sufficient to answer key questions concerning landscape evolution (Williams et al., 2013). This is most readily achieved through the calculation of a DEM-of-Difference or DoD that has applications in both the assessment of process and form (Wheaton et al., 2013) and the assessment of emerging numerical morphodynamic models (Williams et al., 2016b). In its simplest form, a DoD provides a spatially-explicit volumetric assessment of change by subtracting an 'old' DEM from a 'new' more recent later surface (Williams, 2012).

$$\delta E_{x,y} = DEM_{2_{x,y}} - DEM_{1_{x,y}} \quad [2.8]$$

where δE = change in elevation (m), and DEM_2 and DEM_1 correspond to the elevation or z-value of the respective DEMs (2 = new, 1 = old), calculated on a pixel-by-pixel basis as denoted by the subscripts x, y. This result can be statistically aggregated into volumes and areas of change (vertically positive and negative) and in map form as shown below in Figure 2.8.

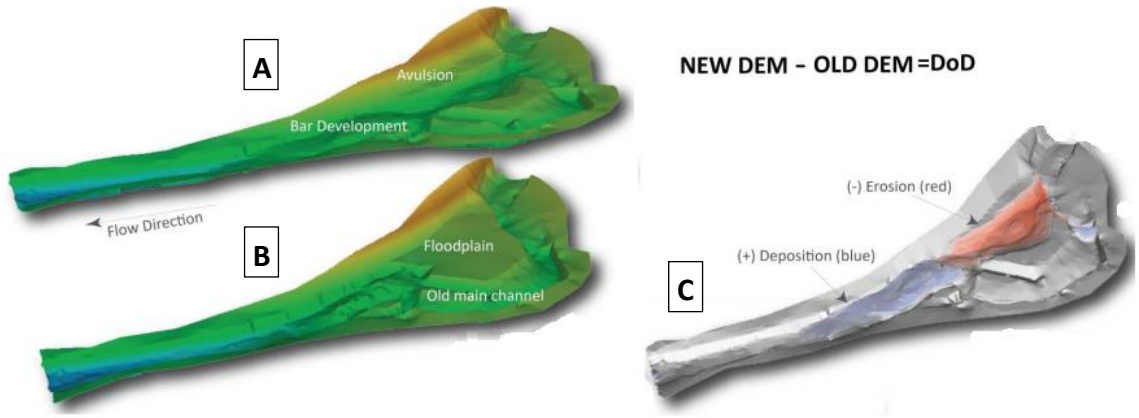

Figure 2.8 Example of the DoD process, whereby the subtraction of the old DEM (a) from the new DEM (b) results in a DoD (c) which when thresholded to account for uncertainty highlights regions of erosion and deposition (figure source, Wheaton et al., 2010).

In this form, negative values imply lowering of the surface and vice versa. Caution should be taken to correlate these directly to a given process, i.e., negative changes imply erosion or subsidence, due to the possibility that unfiltered changes are data artefacts.

2.6.3. Methodologies of error analysis

The ability to infer the distribution and magnitude of geomorphological processes derived from SfM data products is ultimately limited by the understanding of uncertainties within the pipeline process (James et al., 2017). When undertaking a DEM-DEM comparison, the error inherent in both surveys renders volumetric errors doubly sensitive to DEM quality (Brasington et al., 2003). The GCD framework detailed by Brasington et al. (2003) and refined by Wheaton et al. (2010) and Vericat et al. (2017) describes a process for analysing geomorphological change as a three step process, involving: (i) individual assessment of uncertainty within each DEM; (ii) calculation of the propagation error associated with subtraction uncertain DEMs; and finally (iii) assessing the significance of observed changes. Step two in the process may be supported by the use of the standard theory of errors (Taylor, 1982). This states that for mathematical operations involving the calculation of a derived variable, for convention, U, by subtraction or addition of two further variables (z_1 and z_2) and their associated error terms δz_1 and δz_2, the total error in u can be estimated as the sum of the two component errors in quadrature:

$$\delta u = \sqrt{\delta z_1^2 - \delta z_2^2} \quad [2.9]$$

where, δu is the propagated error (here, in m). This approach has been used by Brasington et al. (2003) and Lane et al. (2003) to identify a minLoD, in which the standard deviation of error (SDE, based on independent check data), which represents the 1-sigma or 0.68 confidence interval for observed changes occurring by chance sampling error. This can be extended to place the analysis into a flexible probabilistic framework, by representing the distributing the propagated error as a Student t-statistic;

$$\delta u_{DoD} = t\sqrt{SDE_1^2 + SDE_2^2} \quad [2.10]$$

where SDE_1 and SDE_2 correspond to the SDE derived from both the initial and repeat survey and the t-critical value represents a probabilistic threshold, calculated from the following equation:

$$t = \frac{|Z_2 - Z_1|}{\delta u} \quad [2.11]$$

and in which $|Z_2 - Z_1|$ is the absolute value of the DoD (though signed values can also be used and by implication identify the direction of change). In this form, a level of detection can be set at any significance level based on the concern for the incorporation of Type I and Type II (errors of commission and omission). In this context, Lane et al. (2003) used a 68% confidence interval, where changes are treated as significant where t is equal to or greater than unity, whereas Wheaton et al. (2010; 2013) used the more conventional 95% significance level. Implementing a probabilistic minLoD in this manner assumes the errors to be random, independent and follow a normal distribution. Such assumptions are likely to be violated, in particular due to spatial autocorrelation of systematic errors, however, the approach does allow for the representation of DEM error as a spatial phenomenon. In its simplest form, this can be achieved by classifying DEM errors by survey type or surface cover (e.g., Brasington et al., 2003) or through a complex modelling process, such as a fuzzy inference system (see Wheaton et al., 2010).

Whatever approach is used, such modelling is likely to require some quantification of surface error, which as discussed extensively above in section 2.5 is a non-trivial problem.

2.6.4. **The Management of 'Big' data**

The emergence of hyperscale survey methods data has resulted in the production of point clouds of increasing density and spatial extent (Vericat et al., 2014), which increasingly presents a computational challenge to store, visualize and analyse. To date, most approaches to quantifying change detection typically rely on the comparison of raster surface models as described above. Such analysis requires the conversion of 3D point cloud data into raster DEMs involving some approach to numerical interpolation (Lague et al., 2013). To manage the increasing density of information used in the creation of surface models, spatial decimation and filtering of the raw point cloud is typically applied to simplify this process. A popular framework to achieve this is through the Topographic Point Cloud Analysis Toolkit, ToPCAT after Brasington et al., (2012). This algorithm seeks to achieve an intelligent decimation of large point clouds in a computationally efficient manner, that gives rise to gridded or voxelated data structures, populated with a statistical description of the subgrid elevation distribution. The principles underlying this approach are shown in Figure 2.9 (Brasington et al., 2012). In its 2D form, a numerical grid is used to segment the point cloud at a user defined resolution, and within in each cell, the statistical moments of the elevation distribution are obtained. These observations are then used to provide a framework to create a low-frequency tessellation of the surface based on the mean, minimum or maximum elevation and this surface is used to detrend the point cloud locally. The goal here is to distinguish sub-grid variability that reflects variation around a broader scale trend (i.e., the topography) defined at the frequency of the numerical grid. This approach has been used successfully to provide both metrics of surface uncertainty but also to relate this estimate to sub-grid phenomena such as particle size (Williams et al., 2011, 2013; Javernick et al., 2014; Vericat et al., 2014; Smith and Vericat, 2015). The ToPCAT algorithm is incorporated into the GCD tool, providing a simplified workflow to store, load, and analyse hyperscale data within a conventional GIS environment.

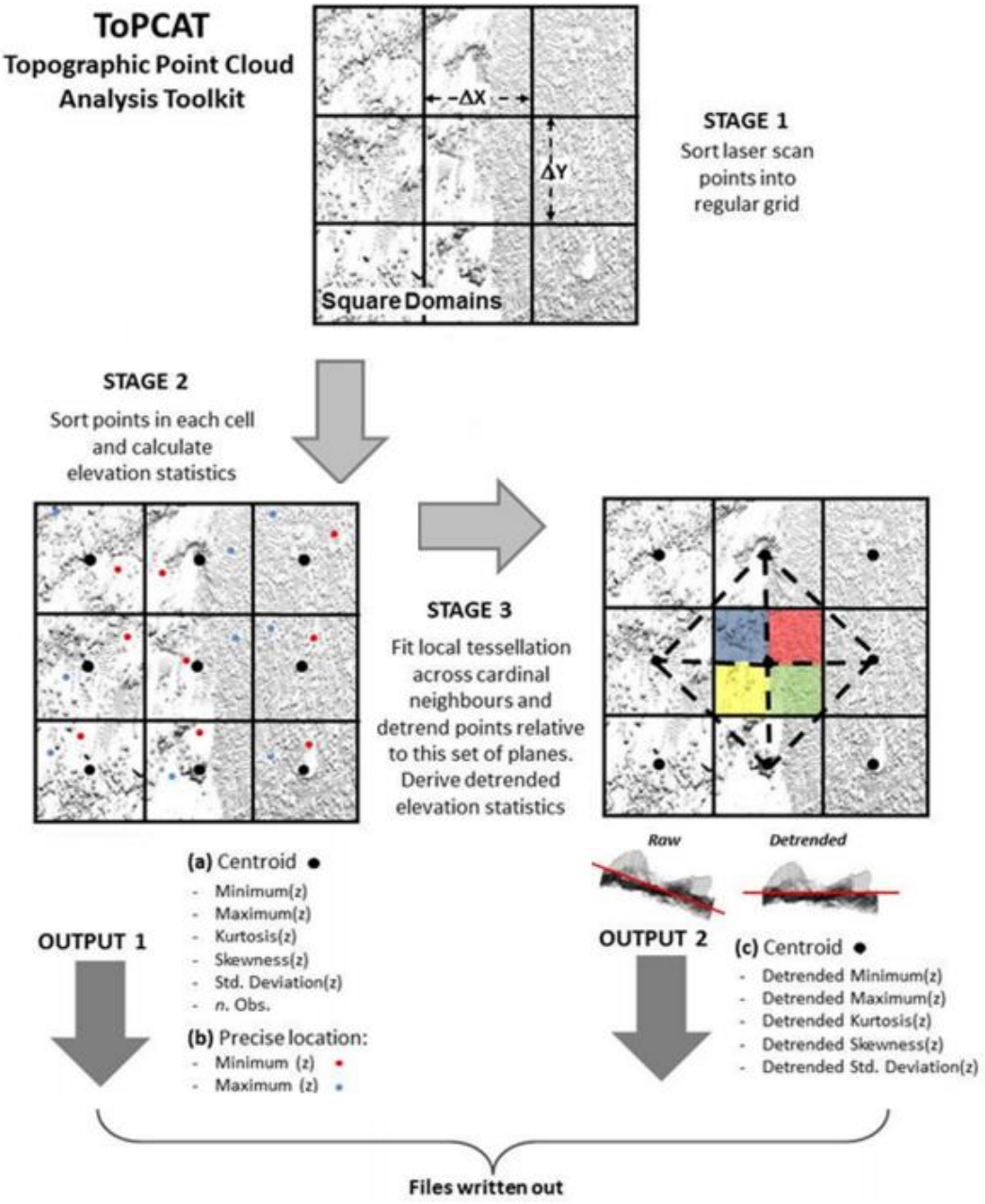

Figure 2.9 Schematic representation of the three stages of the ToPCAT point cloud analysis kit. A user defined grid is specified to segment the point cloud and sort into cells, from which elevation statistics are derived, before fitting local tessellation to derive detrended elevation statistics (figure source, Brasington et al., 2012; p. 6).

2.6.5. Examples of DoD use within fluvial morphology

DEMs of Difference have been applied to a wide range of problems in fluvial geomorphology over the last two decades, and using a wide range of geomatics data sources, including: RTK-GNSS survey (Brasington, 2000); aerial photogrammetry (Brasington et al., 2003; Lane et al., 2003); TLS, (Milan et al., 2007; Wheaton et al., 2010;

Williams et al., 2011; Wheaton et al., 2013); and Structure-from-Motion (Marteau et al., 2017). These studies can broadly be categorized by their objective, to either map spatially regions of change and by inference bed mobility and turnover, or to provide quantitative estimates of bed-load transport and process mechanisms (Brasington et al., 2003; Wheaton et al., 2013). For example, using DoDs of the braided River Feshie, Wheaton et al. (2013) attempted to quantify the relative contribution of key braiding mechanisms to channel change based on the original perceptual model of braided rivers by Ashmore (1991). In contrast, in the case of the latter, the morphometric approach enables an indirect alternative to the difficult task of directly sampling bed load rates as reviewed in detail by Vericat et al. (2017). A related approach to this is described by Redolfi et al. (2016) who use DoDs to determine a representative particle transport length or step-length, from which estimates of sediment transport can also be inferred.

The estimation of change detection through DoDs represents a convenient 2.5D simplification that exploits the speed of raster processing and assumes that the principal vector of topographic change is vertical. However, the conversion of raw point cloud to decimated grids in this process inevitably results in the loss of more complex patterns of change, such as bank retreat, hollowing out and overhang collapse that can ultimately only be represented in 3D (Brasington et al., 2012; Brodu and Lague, 2012).

2.6.6. Comparison of Point Clouds based approaches to change detection

A DoD analysis provides a 2.5D analysis of change, however, for the purposes of detecting 3D change, analysis through the use of point clouds is required. Point clouds remain the most primitive, yet fundamental form of surface representation, with the potential to outperform model-to-model methodologies (Memoli and Sapiro, 2005; Kim et al., 2014). Cloud-to-Cloud (C2C) approaches directly compare two point clouds and are the simplest and fastest type of technique, specifically designed to identify rapid change between dense point clouds (Barnhart and Crosby, 2013; Brodu and Lague., 2012). An alternative approach, of Cloud-to-Model (C2 m) analysis is perhaps the most common technique, comparing point clouds to a refined triangulated mesh, although uncertainties remain difficult to quantify (Lague et al., 2013). Despite previous uses, these two approaches remain unable to provide explicit representation of data

uncertainties that account for instrument and registration errors (Brodu and Lague, 2012).

In response, Lague et al. (2013) developed the Multiscale Model-to-Model Cloud Comparison (M3C2) algorithm, specifically designed for application to fluvial terrain datasets. This algorithm detects change with minimal cleaning and processing of scan data (Barnhart and Crosby, 2013; Earlie et al., 2013; Lague et al., 2013). The M3C2 algorithm employs a two stage process based on the calculation of point normals (the local orientation of the point cloud) and difference computation, see Figure 2.10 below. The approach uses the concept of 'core points'; a subset of user defined points to reduce computational intensity (Barnhart and Crosby, 2013). The point cloud P_1 may then be reduced to P_{1c}, however, during analysis, any point in P_1 within the vicinity of *i* within P_{1c} is retained for analysis. For a given core point, *i*, a normal vector is defined for each cloud by fitting a plane locally to the points within a user-defined radius, *D/2*. A standard deviation between neighbours and best fit plane is used to calculate surface roughness. The choice of normal is dependent upon the user – conventionally the orientation of the *a priori* cloud is used to define the direction of change (Brodu and Lague, 2012; Lague et al., 2013). Following point normal calculation, a cylinder normal to point *i* is projected across the two point clouds, with a user-defined radius. All points residing within the cylinder are then spatially averaged to compute mean surface positions and their respective difference.

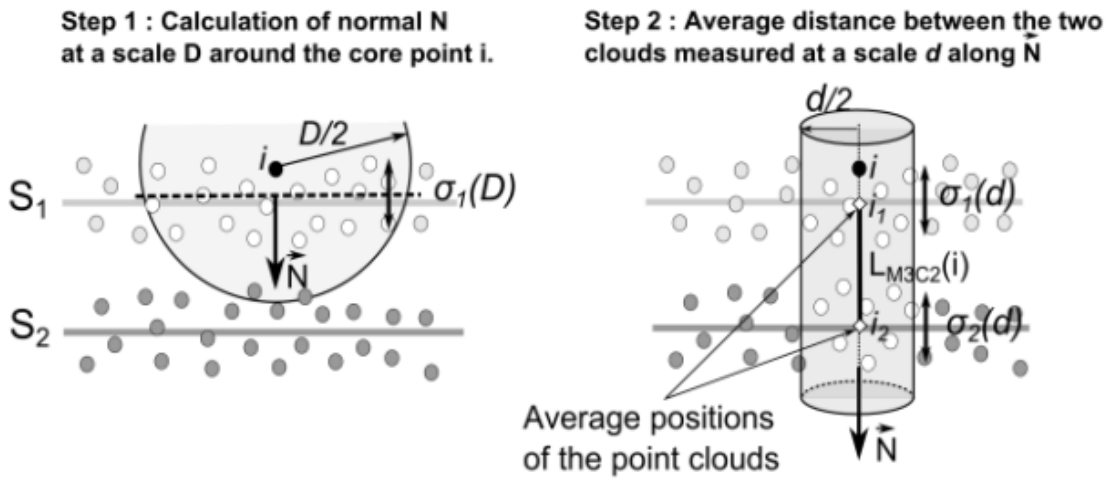

Figure 2.10 Graphical representation of two stages of the M3C2 algorithm applied to two planar surfaces; point normal calculation and computation of difference *N*. The radius *D* is user defined dependent on survey requirements and computation limitations. The subsequent projection of a cylinder, radius *D/2* between clouds from which points are spatially averaged to compute the mean surface (figure source, Lague et al., 2013; p. 7).

Error is accounted for via a spatially variable confidence interval (SVCI), which accounts for the impact of local point density and surface roughness upon precision. The SVCI limits may also be parameterized using the additive root mean square (RMS) error registration error (where available, for example in a TLS survey). The resulting confidence intervals provide a conservative approach to the detection of changes in comparison to C2M, with many displacements found to be insignificant (Barnhart and Crosby, 2013; Lague et al., 2013). Utilising M3C2, Lague et al. (2013) were able to achieve detection of change (95% confidence interval) down to 5 mm along a meandering bedrock river. Whilst the M3C2 method would provide for a truly 3D analysis of geomorphological change, it is designed primarily to handle laser ranging data and more importantly – for the purposes of managing 'big' data – the raw SfM cloud at the catchment scale would be computationally onerous.

2.6.7. **Conclusions**

The reliable analysis of change detection remains crucially dependent on the ratio of the magnitude of geomorphological change to uncertainty in the input datasets (Brasington et al., 2000; Wheaton et al., 2010; James et al., 2012). Given the typically shallow magnitude of erosion and deposition in fluvial systems, and the challenges of accurate terrain survey and modelling, the assessment of geomorphological change in rivers is

likely to remain a persistent area of interest. The recent emergence of the point cloud based approaches to the problem, such as the M3C2 algorithm, provide a means to use raw survey data in their most fundamental form, without incorporating global interpolation uncertainty (Lague et al., 2013). While this approach appears encouraging, there is a computational attraction of raster modelling scales of the areal extent to be examined. In this context, catchment scale SfM studies are thus likely to benefit from data efficient and computationally fast raster DEMs and DoDs. A major drawback remains the lack of error modelling studies utilising SfM generated point clouds – with the majority of techniques developed (Brasington et al., 2003; Milan et al, 2007; Wheaton et al., 2010) based on traditional photogrammetry, TLS and LiDAR derived point clouds. James et al. (2017), provide a SfM specific error model based on a Monte Carlo analysis of the bundle adjustment that is used to determine tiepoint uncertainty across 98 images within the sparse cloud. Again, while such an approach is attractive, the sheer number of photographs required to model large areal extents are likely to render this approach impractical for this study.

2.7. Opportunities for further study

Recent geomatics developments are creating the opportunity for SfM to undertake temporally and spatially dense 4D analysis of the Earth's surface topography (Smith et al., 2016). Whilst broad scale applications of SfM are beginning to emerge, there remains a key research gap to develop and test strategies to optimise the SfM workflows and examine the reliability of the generated data products to quantify geomorphological change accurately and robustly.

These specific gaps are directly related to the research aims stated in Chapter 1. First, these focus on the development of an 'acquisition-to-data product' workflow for 3D image-based reconstruction of large fluvial systems. This review has clearly identified opportunities to explore different camera network geometries and examine the use of multiple camera angles to minimize systematic errors that have been observed to arise in the self-calibrating bundle adjustment. This approach has the potential to relax the demands for a dense control network that is likely to be difficult to establish over a large region of challenging and inhospitable terrain. An additional area to consider, is to

maintain comparable densities of imagery to those used in standard UAS studies, rather than the lower densities used by, for example, Dietrich (2016). A key challenge here will be to explore whether it is possible to develop strategies to build such dense image blocks over wide areas without the need to decompose the region into overlapping blocks that are complex to merge. These demands pose added concerns for both field data collection and in terms of subsequent computational processing. The processing times expected from large imaging sets are significant, however, there exists potential when restricting the SfM algorithm to regional searches for image pairs.

The second research aim relates to the development of a quantitative error model to support SfM modelling of fluvial systems. Current models have tended to address uncertainties and errors within point clouds derived from laser ranging techniques. These are unrepresentative of the most significant errors within the SfM pipeline. Whilst there are emerging methods to derive tiepoint precision within SfM, the scale of systematic errors and numbers of imagery may nullify the use of a similar approach to James et al. (2017). There is a clear need, through the use of repeat surveys, to begin to quantify these errors across multi-temporal large scale SfM studies.

3. Chapter 3: Study Site and Data Acquisition

Abstract

The Dart River, New Zealand, provides an exciting opportunity to derive catchment scale terrain models to assess channel evolution in the aftermath of a significant external sediment input. The varied landscape of the Dart, however, presents a number of photogrammetric and logistical challenges, which are covered in detail. This chapter subsequently addresses the use of black box software Agisoft Photoscan and how the key processing steps taken relate to the fundamentals of SfM discussed previously. The remainder of the chapter details the collection of control datasets and the error metrics presented during the results of this thesis.

3.1. Chapter content

Following from the review presented in Chapter 2 the challenges of developing a broad-area, catchment-scale SfM survey of the Dart River that extends over >60 km^2 of remote and hostile terrain can usefully be decomposed into issues relating to (a) data collection and (b) data processing. The key aspects of this classification are described in Table 3.1 which highlights the implications for SfM modelling.

Table 3.1 Data collection and processing challenges faced when implementing large scale studies.

Source	Logistical Challenge	Photogrammetric compromise
Data collection	Low density of ground control; sub-optimal distribution of ground control; long flight times; high flight altitudes and speeds; variable geomorphological characteristics.	Sparse, sub-optimal network of exterior constraints; variable illumination due to changing solar angle; reduced keypoint fidelity and resolution; local variations impacting a global solution.
Data processing	Large volume of images to cover area; doubling image numbers if applied with a multi-view camera system to provide convergent image network.	Survey extent necessitates limited side and endlap; sub-optimal settings used within APS; reliance on derived camera orientation; reduction in keypoint quality with the up-take of compressed imaging formats.

The methods detailed in both this Chapter and Chapters 4, 5 and 6 attempt to mitigate these challenges and assess their impact and practicability. This specific chapter aims to provide an overview of common elements that pertain to all later aspects of the research, in essence providing a generic overview of some of the key methods used and data collected. This includes:

(i) an overview of the study area, the Dart River, focusing directly on localized photogrammetric concerns associated with modelling this system;

(ii) a description of the approach to the acquisition of key data for the SfM modelling is then provided (note that specific modifications to the data collection strategy are dealt with in the key results chapters 4, 5 and 6).

(iii) following on from the theoretical discussion of SfM in Chapter 2, the specific processes and workflows as used to implement SfM in Agisoft Photoscan (APS) are described in detail.

(iv) the acquisition of additional independent control data sets for model evaluation is discussed;

(v) a review of how dense point clouds are processed to generate raster surface models; and finally

(vi) the error metrics used throughout this thesis are defined and discussed.

3.2. The Study Site

3.2.1. The Dart River Catchment

The Dart River drains a 630 km^2 glaciated catchment in the Southern Alps of Western Otago, on the South Island of New Zealand (McColl and Davies, 2011). The catchment is bounded by the Humboldt Mountains to the West and the Barrier and Snowdrift ranges to the North, with watershed peaks that range between 1300 – 2800 m. The river is glacially-fed by the now small (c. 7 km^2) Dart Glacier. The Dart River stretches over 60 km from its source to its mouth in a large 2.5 km wide delta at the head of Lake Wakatipu (Figure 3.1). At the regional Last Glacial Maximum extent (c. 25,000 BP) the Dart Glacier was part of a much larger ice cap that filled the Wakatipu basin and terminated some 135 km downstream at the Southern end of Lake Wakatipu at Kingston. The country rock is highly variable throughout the catchment, but dominated by green and grey schists, but also includes outcrops of pounamu (Māori), a highly-prized nephrite jade treated as sacred by the local Māori tribes (iwi).

The Dart is a wide, actively braided gravel bed river that varies in width between 0.6 – 2 km. The channel is constricted between high terraces at three distinct locations and becomes single thread within a narrow confined 'gorge' in the headwaters. The river is gauged by the Otago Regional Council at the Hillocks recording station, at river km 12 (catchment area, 563 km^2) associated with a narrowing and vehicle crossing. The observed flow record began in 1996, and the mean annual flood is estimated at c. 900 $m^3\ s^{-1}$, although peaks exceeding 1450 $m^3\ s^{-1}$ were recorded in 1999, 2013 and 2015. The median discharge is 51.49 $m^3\ s^{-1}$.

Average channel bed slope along the mainstem is 0.003 (m/m) and surface grain-size declines downstream rapidly from D_{50} ~0.14 m upstream of the Slipstream landslide to 0.013 m above the delta (Wild, 2012).

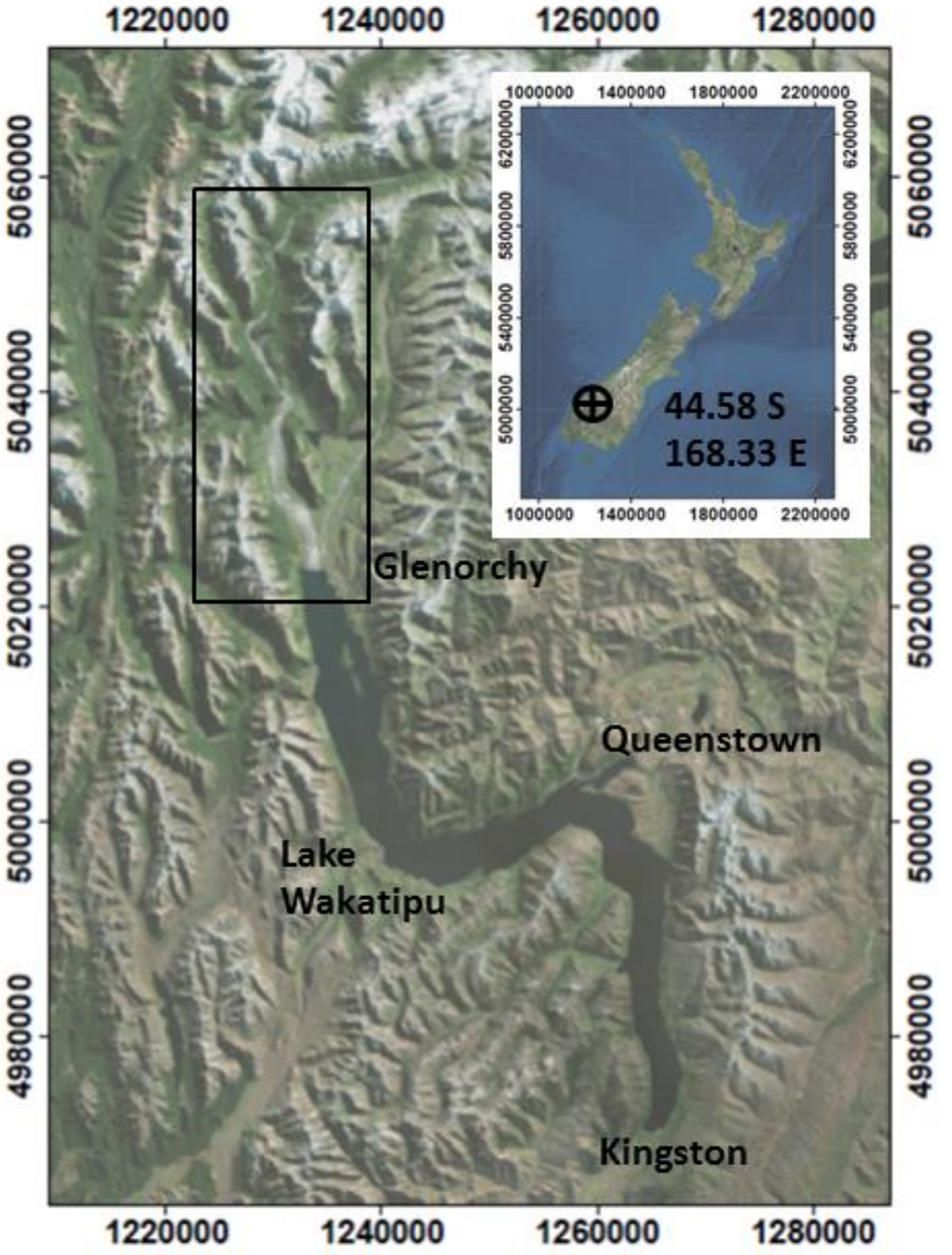

Figure 3.1 Aerial Imagery of the Dart catchment in proximity to the head of Lake Wakatipu. *Projection: NZTM 2000*

Recent estimates suggest annual sediment delivery from the Dart to Lake Wakatipu has averaged ~0.27 x 10^6 m^3 yr^{-1} over the last 20 years, with significant progradation of the delta during this period (Wild, 2012). There is regionally high seismicity in the region, with a large number of active faults nearby including the major dextral-transpressional Alpine fault that demarcates the boundary between the Australian and Pacific plates. This fault is thought to generate earthquakes with a moment magnitude of M_w >8 with a return period of 250-500 years. The last major rupture is dated as 1717 (Sutherland et al., 2006). Landsliding occurs both co-seismically and aseismically and has played a dominant role shaping the catchment at a range of scales. There is evidence of multiple major slope failures within the catchment, including with evidence of past rock avalanche events that are likely to have dammed drainage in the past (McColl and Davies, 2011; Wild, 2012).

Figure 3.2 Both images taken two days apart from opposing sides of the Bridge at the Hillocks, (A) a low flow of ~50 m^3s^{-1} and (B) a peak flow of ~ 700 m^3s^{-1}.

The catchment cover comprises a mix of privately owned (and Crown-leased) grazing land on the lower slopes and valley floor, with extensive Southern Beech (*Notofagus* spp.) forest at higher elevations and Alpine conditions above 1700 m. Approximately

70% of the catchment lies within the state owned and managed Mount Aspiring National Park. A number of popular hiking trails exist throughout the catchment and are maintained by the Department of Conservation (DoC). The most famous of these is the multi-day Dart-Rees track that attracts over 10,000 visits per year. The township of Glenorchy, which sits at head of Lake Wakatipu, on true left of the Rees-Dart Delta, is a major centre of tourism with extensive accommodation and tour-based services, including a major company providing jet boat and canoeing activities along the Dart River. There is significant local concern over elevated risks of flooding associated with sedimentation and progradation of the delta following the recent landslide at Slipstream.

3.2.2. Catchment Regions

To support description and the explanantion that follows in subsequent chapters, the study region is here compartmentalised into a series of distinct regions as shown in Figure 3.3. These represent morphologically distinctive reaches that raise their own unique challenges for an SfM-based survey strategy, beyond the broader issues of scale and access that pertain to most areas.

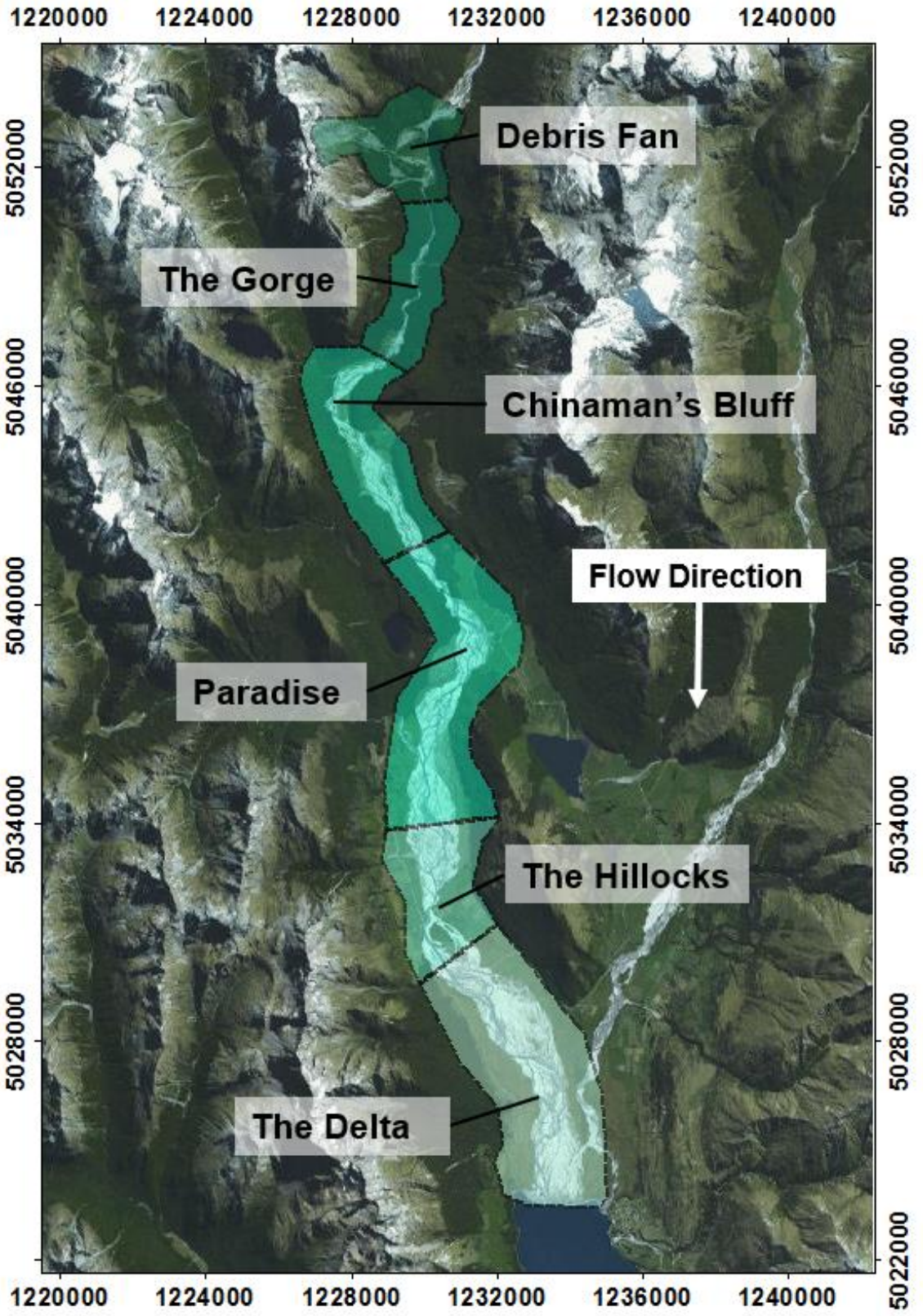

Figure 3.3 The reach-by-reach classification of the Dart River. These specific reaches will be referenced throughout this thesis and relate to the major landmark/features of the associated areas. *Projection: NZTM 2000*

3.2.2.1. Slipstream Landslide and Debris Fan

Following sustained heavy rainfall on the 4/1/2014, a major landslide occurred at Slipstream on the true right bank of the Dart. The event was first noticed by the jet boat operators (Dart River Safaris) who observed a sudden fall in the river level and raised the attention of the Department of Conservation office in Glenorchy. Subsequent helicopter reconnaissance by a science team from GNS Science (S. Cox, pers. comm.) indicated that the landslide had extended across the existing fan surface and blocked the Dart River creating a large in-line lake upstream, inundating Daley's Flat (see Figure 3.4 below). The volume of the lake was subsequently surveyed as an allied part of this research project and found to contain c. 5.8 M m^3 water with depths locally exceeding 12 m. Modelling suggested that given the low crest height of the dam, which partially breached shortly after the event, there was no immediate risk of catastrophic failure and downstream flooding. Following the partial breach, much of the toe of the fan, incorporating sediment from this and recent debris flows has started to be reworked and transferred downstream. There is an intensively braided fan immediately downstream of the dam crest through which the Dart now flows before it is constrained by valley terraces.

Figure 3.4 (1) Photograph obtained of the Slipstream debris fan in January 2014 and (2) the lake that formed in the aftermath of the damming of the Dart River (Cox et al., 2014).

The fan surface, largely vegetated before landslide activity first began in 2009, has since been covered by debris comprising a range of clast sizes, from large boulders to fine silts (Cox et al., 2014). This surface is an ideal target for photogrammetric reconstruction, with an accessible (if soft) terrain to distribute control targets, textured cover that

generates unique shadows and has only limited vegetation. The surrounding steep slopes (including the landslide gully) might be expected to present some problems due to difficulty of placing control across an effective vertical range (providing a robust 3D distribution) and due to terrain shadowing and the mix of spectrally dark forest and bright exposed sediments. As described below in section 3.4, this site was also surveyed at the time of photo acquisition in 2014 using long-range Reigl VZ-1000 TLS to provide spatially dense independent check data to assess the photogrammetric reconstruction.

3.2.2.2. The 'Gorge'

The reach downstream of Slipstream is incised into high Quaternary terrace deposits and comprises a comparatively narrow, single thread channel varying in width from 40-100 m. Nicknamed 'the Gorge', this reach is approximately 5 km long. The surrounding terrace and lower slopes of the catchment are densely forested, while the channel is tightly confined, with only one area with significant floodplain development on a tributary fan. The active channel width is c. 40% inundated at low flow with the remaining width comprised of well-imbricated lateral and a few mid-channel gravel bars. There are a number of large boulders in the upper part of the reach, known locally as the Cathedral rocks and thought to be runout from a major early Holocene rock avalanche (Turnbull, 2000). These boulders create sedimentation zones in their wake leading to bar development and may themselves protrude 4-6 m above the water level.

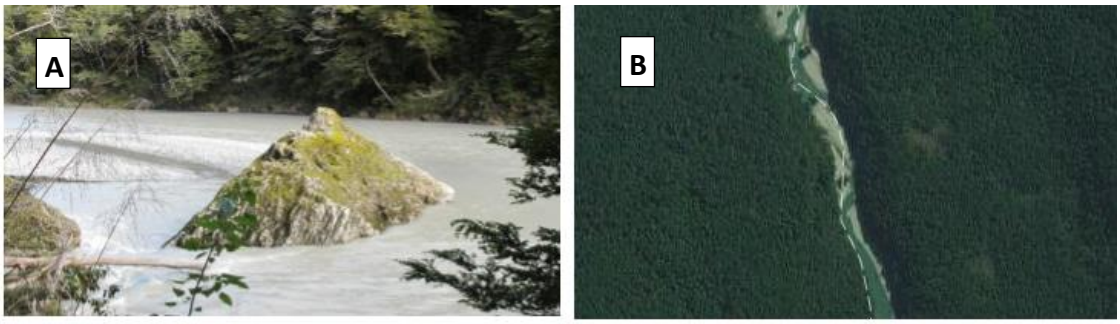

Figure 3.5 A) Significant 'Cathedral' rock deposit within the Gorge (Brasington et al., 2014); and B) Aerial imagery of the surrounding vegetation density (Google Earth, 2017).

This region presents significant challenges for a SfM workflow, as the surface type and ground control are far from ideal. The single thread nature of the channel presents relatively few stable bars suitable for GCP deployment and many of these are: a) difficult to access, largely reachable only by jet boat; and b) have a high risk of inundation

between ground deployment and photo acquisition. Much of the active channel is also shaded by riparian forest cover that obscures any bank positioned targets. As much as 40% of the active channel is inundated at low flow, and with rough flow there is a high potential for spurious keypoint detection and correlation. Furthermore, the active gravel contrasts strongly with the dense vegetation cover that will comprise the majority of any photographs acquired. This will lead to a challenging photographic context, likely to lead to over or underexposure of large areas of the scene and further potential for poor keypoint detection and matching.

3.2.2.3. Chinaman's Bluff

The Dart emerges from the 'gorge' at Chinaman's Bluff, which is also the roadhead of the gravel track on the true left of the Dart. The bluff is a glacially truncated cliff with a peak of 927 m, and gains > 200 m in elevation abruptly at the true left channel margin.

Figure 3.6 Chinaman's Bluff on the right, viewed immediately downstream looking North, with the Dart emerging out of the gorge region in the distance.

The channel widens dramatically within this reach, with the active channel belt reaching 0.9 km, characterized by a labile braided morphology with little floodplain. While the channel bed is largely free from vegetation cover, the widespread and complex pattern of inundated anabranches, even at low flow, will present the potential for frequent spurious keypoint detection and matching. Access to deploy ground control is limited to a small number of access points on both the true left and true right and again the potential for disturbance during moderate to high flow events is high. There is limited scope for easily visible high spots above potential high water levels to position GCPs.

3.2.2.4. Paradise

The region named Paradise (named after the local station or farm) marks the reach between Mount Earnslaw and Mount Alfred. The reach incorporates a makeshift jetty maintained by Dart River Safaris and is accessible by car at its upper and lower end.

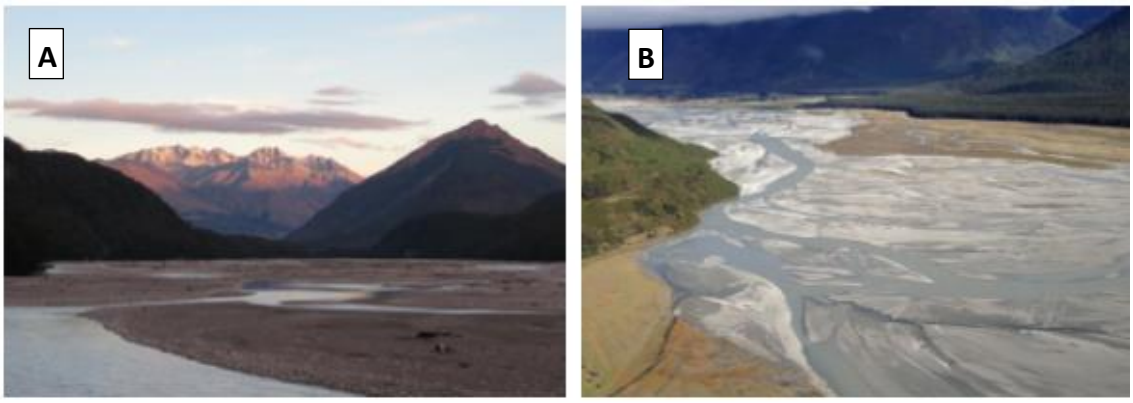

Figure 3.7 (A) The view upstream from Paradise and (B) the extensive braidplain downstream from Paradise (Brasington et al., 2014).

Here, the braidplain reaches a maximum width of 1.5 km in a relatively unconstrained valley floor setting. There is ready access to both sides of the channel to deploy ground control; via the Routerburn track on the true right and vehicular access over private land on the true left. Access to the centre of the channel is, however, difficult given the width, depth and flow velocity of the major anabranches which are unsafe to wade. The riparian forest cover is extensive and, while given the channel width, this poses less of an issue for ground control although the image characteristics are likely to generate spurious tiepoints.

3.2.2.5. The Hillocks

Downstream of Paradise, the Dart valley is bisected by the 1375 m Mount Alfred summit, which caused divergence of the historic glacial flows. This point represents a major potential avulsion for the Dart, though recent incision has lowered the true right valley floor and maintained drainage to the east of Mount Alfred. The channel here maintains an extensive braided morphology, before a major constriction at the Hillocks where the river narrows to a single channel of approximately 150 m width. This narrowing facilitates the only road crossing of the Dart, linking the township of Glenorchy to the right bank settlement of Kinloch and the trailhead for the Routeburn track (Figure 3.8).

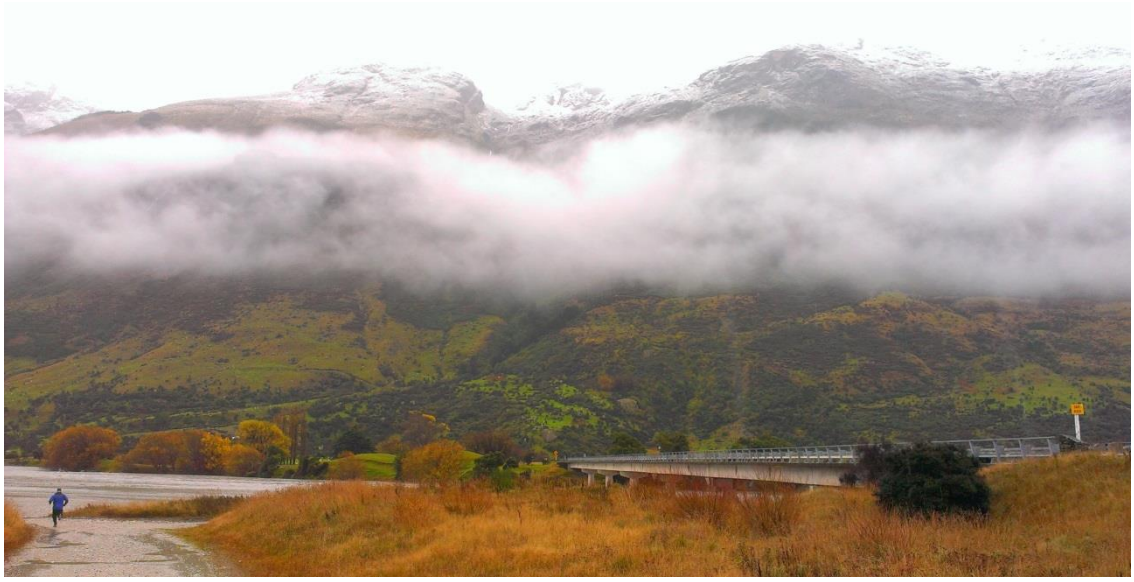

Figure 3.8 The Hillocks, with the hummocky terrain that is characteristic of this region and the presence of the road.

The valley floor at the Hillocks is characterised by hummocky terrain, with small mounds that vary from 4-15 m height and between 10-140 m length. The origin of these landforms has been the source of some debate and has historically been thought to relate to a dissected kame terrace. More recently, McColl and Davies (2011) have proposed the area is actually the runout from a major early-mid Holocene rock avalanche that extends across the full 1.5 km valley floor. The current channel is incised through this deposit and there is no major grade break despite the significant narrowing of the channel through this point. Downstream of the bridge crossing, the channel widens rapidly to a width of c. 1 km and then bifurcates around a major mid-channel island, before maintaining an active width of 1-1.5 km downstream towards the delta. Proximity to the road facilitates placement of ground control in this reach and, given the high banks, GCPs can be deployed at high elevations with low risk of inundation. The relatively narrow width and access to the island in the 2-3 km past the Hillocks provides good coverage of GCPs across the active channel, although the relative high ratio of wet-to-dry bed has implications for the quality of photogrammetric reconstruction.

3.2.2.6. The Delta

Downstream of Mount Alfred, the Dart and Rees river valleys converge to create a broad 3.5 km valley floor. The current active channels are confluent approximately 1.5 km upstream from the current lakeshore, but extensive braided morphologies across the

valley suggest that the two rivers have reworked this area extensively in the past. The Dart is the dominant channel, comprising 90% of the active width. The active braidbelt here varies between 1.5-2.5 km where the main stem and Rees and Dart converge at the delta with Lake Wakatipu. The delta comprises shallow gradient, fine grained (fine gravel with large sandy splays) topsets, and steep forests that dip down to the proximal lake bed at 80-150 m depth. The current lake level averages 310 m ASL which represents the regional base level for the catchment.

There is a moderate true-left to true-right transverse slope across the reach entering the delta, and in the last five years there has been extensive erosion of the true right floodplain (retreating more than 350 m in places). This has caused erosion from major stopbanks and poses a threat to the road servicing the small settlement of Kinloch. Rapid progradation of the delta has also been observed over the last 100 years, with over 120 m of growth calculated from the analysis of historic aerial photographs (Wild, 2012). Such rapid growth poses a major flood hazard for the local communities, most notably the township of Glenorchy which sits at the head of the true left of the lake.

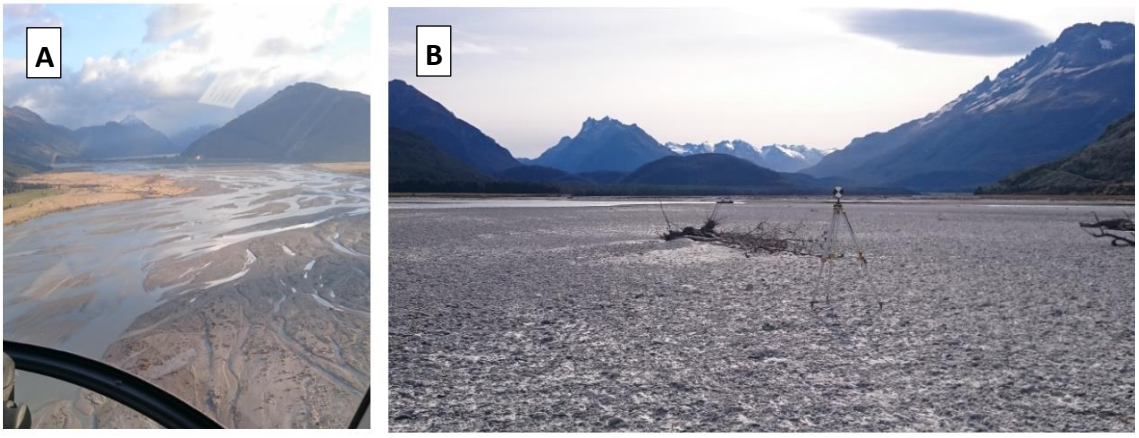

Figure 3.9 (A) Aerial view of the Braiding from a helicopter and (B) Looking upstream from a stretch of floodplain in the Delta region.

This reach is easy to access from both banks, particularly in proximity to Kinloch and Glenorchy. However, the extensive width implies that riparian ground control are set up to 2 km apart with no way of easily accessing the middle of the braidbelt. There is a significant proportion of surface water which once more limits the reliability of

photogrammetric methods. The active channel is, however, largely free from vegetation cover, although established floodplains are typically covered in extensive tussock grass (dominantly *Carex* spp. and *Poa* spp) that can reach 0.5 m high. The wide, flat valley floor limits the effects of shading even at low solar angles.

3.2.2.7. Summary

Throughout the reach there are some common factors that will affect the photogrammetric reconstruction of local areas to a greater or lesser extent. Table 3.2 depicts these key constraints and outlines how they relate to the photogrammetric methods discussed in Chapter 2.

Table 3.2 Localized implications for SfM pipeline respective to each reach identified in Figure 3.3.

Geomorphological Characteristic	Photogrammetric Implication	Regions Impacted
Dense vegetation	Inconsistent scene structure between sampled images due to poor quality keypoints. Result of both homogeneity of surface and tendency for canopies to move.	Gorge, Chinaman's, Paradise, Hillocks, Delta
Wetted channels	Inconsistent scene structure between sampled images due to poor quality keypoints. Result of inability of SfM to accurately model water refraction and movement of surface.	Gorge, Chinaman's, Paradise, Hillocks, Delta
Heavy shading and contrast	Inconsistent scene structure between sampled images due to poor quality keypoints. Result of poor lighting and high contrast in regions of vegetation and schist, compromising keypoint identification.	Debris Fan, Gorge
Minimal exposed proximal areas	Irregular and sub-optimal placement of ground control within the localized area.	Gorge
Valley confinement	Reduction in photo density due to flight line restrictions limiting the number of unique paths.	Gorge

Consideration of the study site characteristics highlights differences in the expected quality of reconstructions should these areas be considered as individual photogrammetric projects (rather than as a whole block bundle adjustment). For example, the Delta region has the potential for a good network of ground control and a range of surfaces that are easily reconstructed in APS. By contrast, the Gorge offers a much more constrained environment for ground control placement and a high probability of spurious keypoint identification and matching due to the extensive wetted width and riparian forest cover. While some of these concerns may be possible to mitigate with extensive local tailoring of the methods to the specific issues, a goal of the research was to reconstruct the whole region. This posed major logistical challenges in terms of the time for the acquisition of ground control, independent check data and the scheduling of photography, as well as the recovery of temporary control marks. Furthermore, due to the topographically inconsistent effect of merging individual photogrammetric blocks (due largely to systematic errors in the periphery of reconstructions), the research aims to reconstruct the entire reach as a single photogrammetric block, inevitably incorporating the local challenges listed above.

3.2.3. Broad-Scale Challenges

The issues facing catchment wide SfM were discussed within Chapter 2 and broadly broken down into data collection and processing. With respect to section 3.2.2 it is clear that the study site contains different regions with specific photogrammetric concerns. It is important to recognize that such local concerns have the potential to propagate into the wider solution. For example, the poor photogrammetric quality of the gorge has the potential to bias the estimation of the exterior and interior camera orientation solved in the bundle adjustment, the effects of which will propagate regionally. Beyond this there are a series of practical limitations that influence the quality of data acquired. The human resources required to distribute and position ground control along a 35 km reach with poor access, limited radio coverage for easy-to-use technologies such as RTK GNSS, are significant and limit the density and optimal placement of control. The larger survey area also necessitates higher flying levels and increased velocities in contrast to traditional UAS surveys. This constraint is not simply economic, as the changing shadow due to solar angle variation throughout the day in mountainous terrain ideally require short surveys close to zenith solar angle. This constraint on image acquisition will serve

to reduce the ground sampling size (to c. 0.05-0.1 m), and require an increase in shutter speed to mitigate motion blur, which in turn will affect optimal exposure. The approaches presented in the following chapter remains subject to these concerns and where possible seeks to mitigate them.

3.3. Data Acquisition

The acquisition of ground-based, aerial and support datasets for this research was closely tied to the objectives of an allied NERC funded project to PI Brasington (NE/G005427/1: *Quantifying the delivery and dispersal of landslide derived sediment to the Dart River, NZ*). This project sought to quantify the morphological, hydraulic and sedimentological response of the Dart River following the major landslide at Slipstream in 2014. The grant was funded under the NERC Urgency scheme (see: http://www.nerc.ac.uk/funding/available/researchgrants/urgency/) and required rapid deployment to the site to capture data in the immediate aftermath of the event. As such, the field data acquisition strategy for the May 2014 survey was undertaken with only limited opportunity for planning and site reconnaissance. Experience gained from this survey and the subsequent data modelling was then used to inform a second repeat survey, one year later in May 2015. As a consequence, the field and aerial surveys undertaken for this research differ in the scope and strategy with the practice in 2015 refined from the lessons learned. The following section details the generic methods used and highlights key strategic differences in two surveys, with implications for the later photogrammetric modelling and data analysis.

3.3.1. Establishing a local GNSS Network

The quality of a photogrammetric survey is ultimately conditioned by the accuracy of the ground control used as exterior constraints. To attain the level of vertical accuracy and precision required to support geomorphological change detection in fluvial systems (ideally 0.1-0.3 m), high quality ground control is required to minimize the propagation of errors (Clapuyt et al., 2016). Given the scale of the study site, using a total station survey to accomplish positioning would be slow and likely to incorporate traverse errors. However, positioning using GNSS observations provides a suitable alternative.

For accurate (0.005-0.05 m), rapid (1-5 s) positioning, two survey methods are available, real-time kinematic or RTK and post-processed kinematic or PPK. In both cases, high positioning performance is established by observing carrier phase measurements and solving the position by incorporating corrections determined from a local, fixed base station sited on an established survey mark or known location. A full description of the theory of satellite based positioning is beyond the scope of this thesis (see for example, van Lekkerkerk, 2017 for further details), so that here only the key practical considerations are outlined.

The two methods (RTK and PPK) use similar algorithms but, in the case of RTK, corrections determined at a base station are transmitted by radio to a roving receiver and accurate positions determined in the field. By contrast, PPK requires logging of all GNSS observables at both the rover and base station and determination of position is calculated after the survey. The use of PPK incorporates a risk that positions could potentially be acquired at times when the remote base station is subject to a poor configuration of satellite constellation, so that accurate positioning cannot later be determined. However, while RTK mitigates such risks, this approach requires the maintenance of VHF radio or GSM (mobile phone) communications, which can be affected by line of sight in complex mountainous terrain. As a consequence, a mix of both RTK and PPK measurements were used to establish control in this research.

Both of the positioning methods described above require data from a base station located at a known point. This can often be achieved by referencing roving observations to a continuously recording GNSS station (CORS) which are maintained at established geodetic marks by national survey agencies, in this case Land Information NZ (LINZ). Accurate differential processing, particularly for short occupation times is, however, only achievable over relatively short (< 1-10 km) baselines due to regional variations in ionospheric delays.

For the study site, three CORS maintained by LINZ and providing GNSS observables in RINEX format, available download within 1 hour, lie within c. 100 km radius of the study site. These include: MAVL at Mavora Lakes (45° 21' 59.46633" S, 168° 07' 05.57202" E), LEXA at Alexandra (45° 13' 51.66242" S, 169° 18' 29.69790" E) and HAAS at Haast (44° 04' 23.53985" S, 168° 47' 07.98976" E). Both HAAS and MAVL were comparable

distances from the centre of the study site (taken as Paradise) at c. 80 km, while LEXA was over 100 km distant. Given these long baselines, a series of local 1st order benchmarks (Figure 3.10) were established using static differential surveying using observations from CORS stations as reference data. A total of four principal local benchmarks were established (Table 3.3) all using occupation times of > 4 hours recording data at 1 s frequency. The local positions were post-processed using data downloaded from MAVL only, as HAAS is on the Australian plate while the study site is located solely on the Pacific plate and the local datum (NZGD2000) is not reliable across this boundary. Data from LEXA were also rejected based on the long baseline to this mark. At each local benchmark, the site was marked using a 50 cm anchored Permamark© and maintained and validated during and after the surveys in both 2014 and 2015.

Table 3.3 Benchmark coordinates established as first order marks based on static differential processing from the MAVL CORS site in 2014 and validated in 2015.

Local Geodetic Mark	Easting (m)	Northing (m)	Elevation (m)
KINL (Kinloch)	1232418.047	5023879.401	314.486
HILL (Hillocks)	1230233.170	5031540.625	338.996
PARD (Paradise)	1231775.679	5039042.591	365.670
CHIN (Chinaman's)	1228380.930	5044694.228	391.253

In addition, a series of second order benchmarks (based on one or more of the marks derived and listed in Table 3.3) were also established as temporary sites to locate base stations in areas likely to be at risk from ground motion, such as on the fan at Slipstream. These second order marks were established during each year and not maintained.

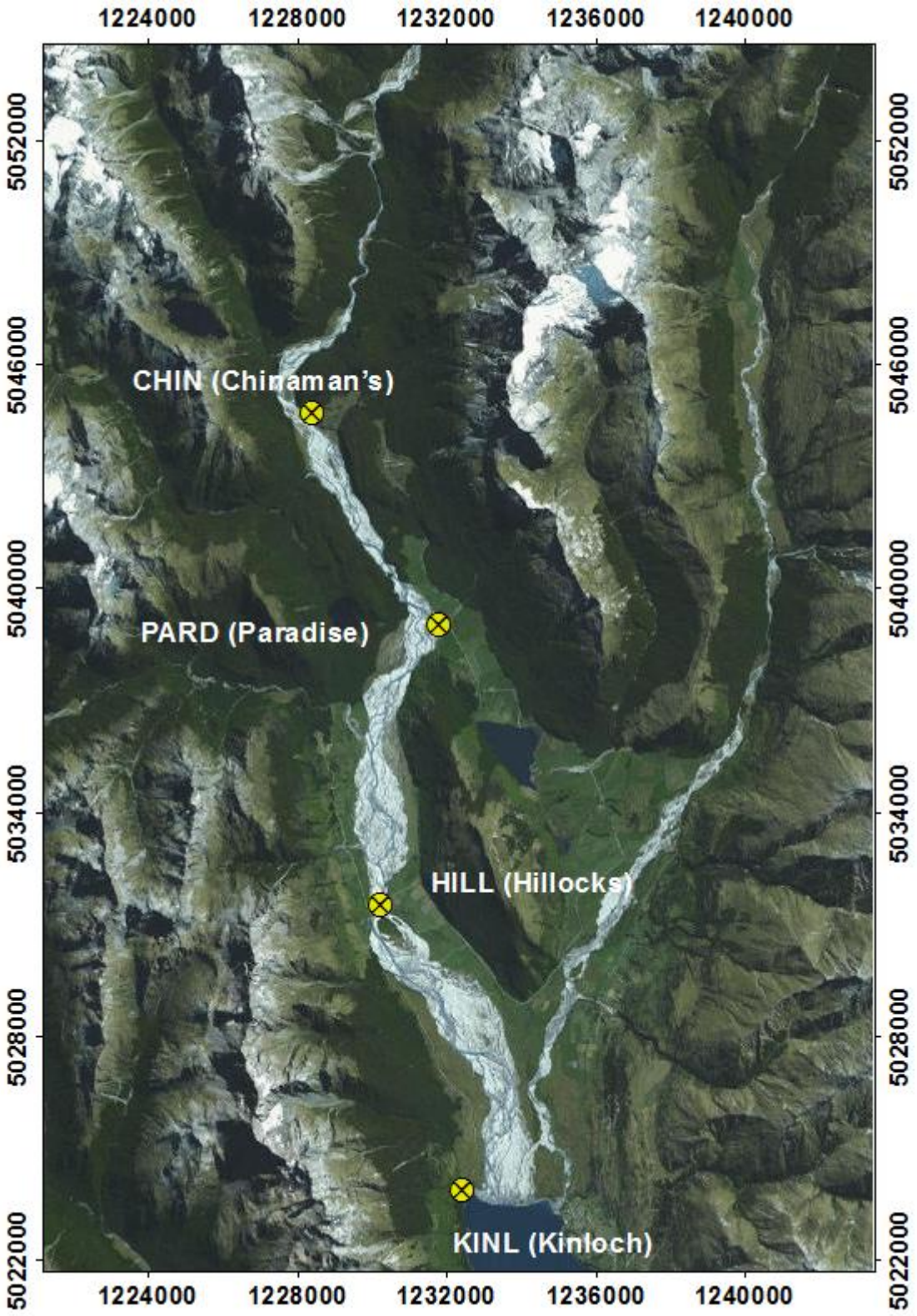

Figure 3.10 Location of 1st order benchmarks established by static GNSS survey based on LINZ CORS at MAVL. All data were processed to provide coordinates in NZGD2000 using the NZTM2000 projection and the NZVD2009 vertical datum (full details of the coordinate system are available at: http://www.linz.govt.nz/information-for/surveyors).

3.3.2. Ground Control Acquisition

For both 2014 and 2015, the surveys were geo-referenced using a network of manually placed GCPs (n = 81 and n = 89, respectively), comprising 1 m^2 heavy-duty, waterproof squares of red tarpaulin pegged to the ground in each corner with 20 cm pegs. The total number of GCPs used is consistent with densities observed in previous large scale surveys (Dietrich, 2016; Immerzeel et al., 2014). In 2014 the centre point of each target was marked by permanent marker to enable the accurate measurement in the field. Analysis of 2014 imagery highlighted a need to delineate the centre point more clearly to ease the process of marker allocation within APS. The subsequent survey in 2015 marked a central black square of 10 cm x 10 cm, reflecting the ground pixel resolution, facilitating identification within APS.

The distribution of GCPs was pre-planned in order to achieve a quasi-uniform density throughout the reach. Locations were selected at approximately 500 m spacing along each bank of the river and these tentative locations marked on a tabled-based GIS overlaid on aerial imagery to support identification in the field. The tablet or a hand-held GPS was then used to navigate to these selected locations, before assessing the suitability of the site directly and, if appropriate, fixing the target in place before surveying the position precisely with GNSS.

Once in the field, however, it became apparent the optimal pattern of deployment was not always possible to achieve due to: a) a requirement to site targets on stable areas that would not be disturbed by rising stage between deployment, survey and collection; b) restricted access to private land; c) difficulty of accessing the location due to changes in the terrain between the aerial image acquisition and time of survey, or dangerous wading conditions; and finally d) overhanging vegetation, particularly in the Chinaman's and Gorge reaches.

As described in section 3.3.1 above, once placed, the targets were surveyed using a combination of RTK and PPK GNSS, with a base station located on a local 1st order benchmark. Observations were made with a range of receivers, including Leica Viva series receivers (GS10, GS15) and a Trimble R8 system. Measurements of targets were made using a hand-held detail pole (Figure 3.11a), and either a timed observation period for PPK measurements (30 s) or precision settings for RTK measurement (0.05 m). All

subsequent GNSS processing, including transformation to NZTM2000/NZVD2009 was completed using Leica GeoOffice.

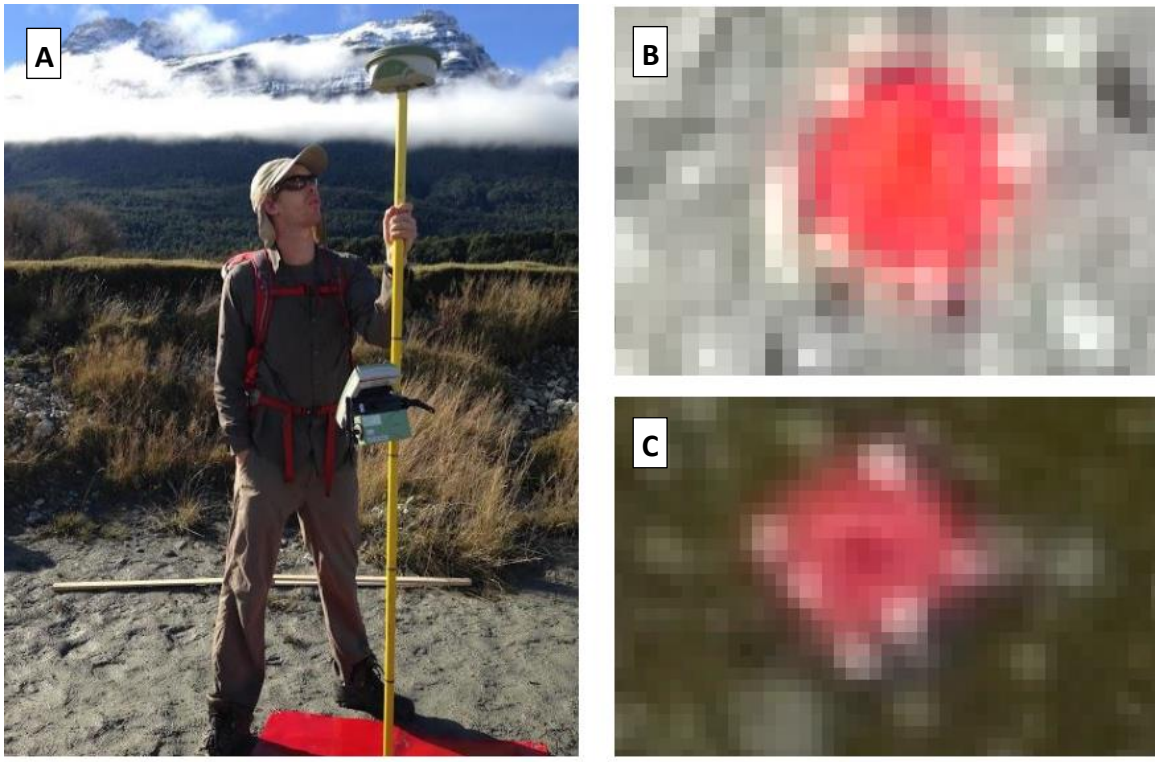

Figure 3.11 (A) Observation collection, with rover mounted to a survey pole and the 1 x 1 m GCP as observed within APS for (B) 2014, with no discernible centre and (C) 2015 GCP, with darkening at the centre of target.

Given the size of the entire study reach, deployment of GCPs took multiple days, even with 2-4 surveyors working in teams. Access to a number of key locations, including the debris fan at Slipstream and large areas of the true right bank upstream of Paradise was only possible using a combination of jet-boat and helicopter support, further constraining logistics. This situation resulted in a significant time elapsing between deployment and image acquisition, in some cases occurring up to a week later. This creates a risk that targets may have been disturbed by wind, water or animals in the interim. Once the survey was completed , each target was checked for visible signs of disturbance and either re-measured or excluded from subsequent analysis. Maps showing the distribution of targets deployed and used either as GCPs or ICPs in both 2014 and 2015 are shown in Figure 3.12.

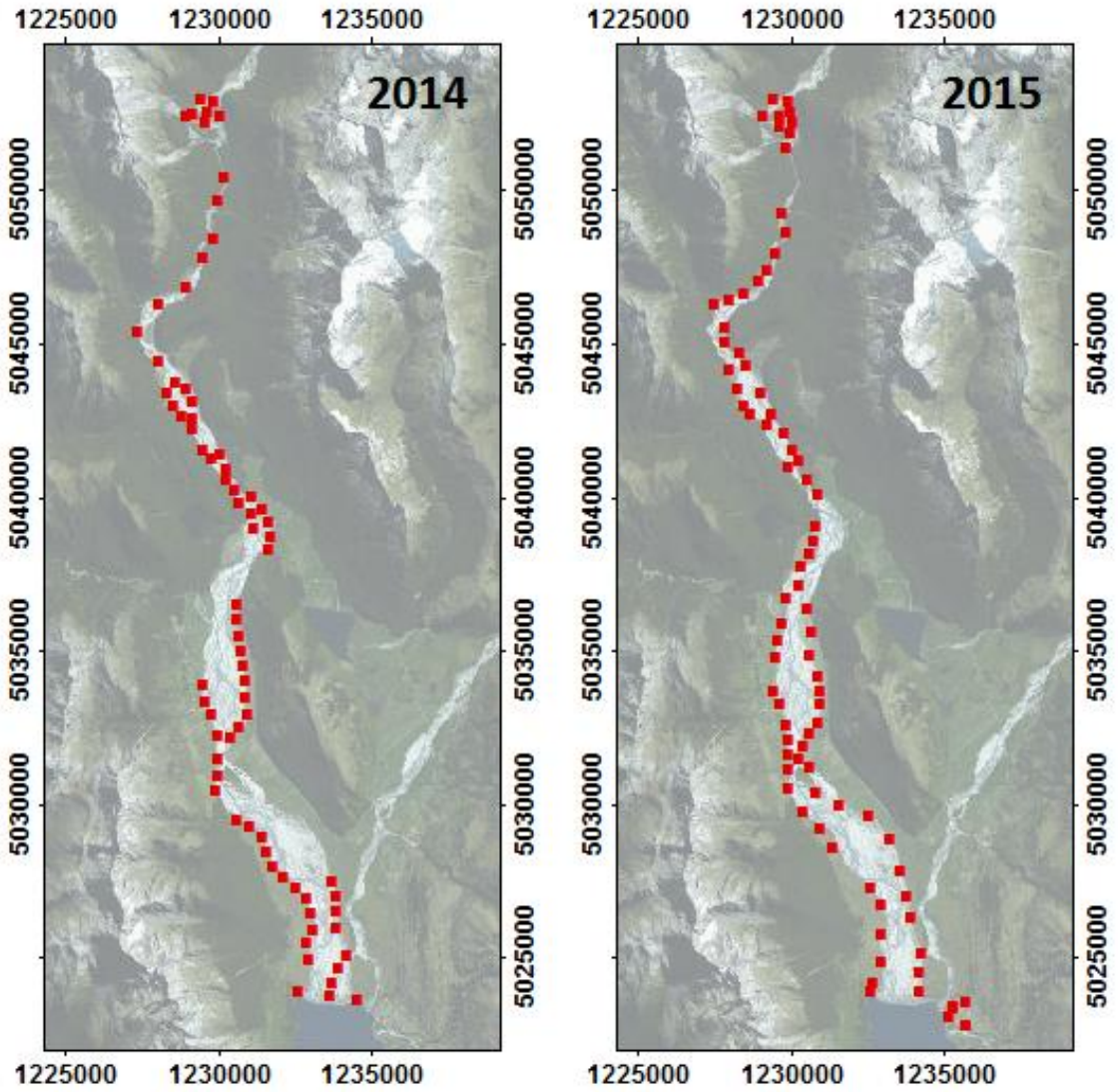

Figure 3.12 GCP distribution for both 2014 and 2015 surveys. The 2014 survey was limited in ground control deployment in the region of the Routeburn and immediately downstream of the Hillocks on the true left. In contrast the 2015 survey was lacking in ground control on the true left by Paradise. *Projection: NZTM 2000*

3.3.3. Image Capture

As described above in section 3.3 the survey strategies in 2014 and 2015 differed, following the revision of approaches due to prior experience and further laboratory modelling and testing of approaches undertaken between the surveys. Therefore, the approach to image collection is described separately below.

3.3.3.1. The 2014 survey

The 2014 survey comprised imagery captured from a nadir orientated Nikon D90 DX SLR camera attached on a rigid mounting attached to the helicopter skid. A fixed 28 mm

lens was attached with the focus set at infinity and exposure settings f/6.3, 1/1000s with images recorded in RAW (NEF) format. The camera sensor has a native resolution of 4288 x 2848 pixels. The settings were tested by performing a short test flight and capturing a series of images to check for exposure and clarity on the day of the survey.

The survey flight was undertaken by Glacier Southern Lakes Helicopters, who assisted in the design of the camera mounting system. A Eurocopter AS350 was used for the flight and in addition to the pilot, two further staff were deployed on board; one to assist with navigation and flight planning; and a second to coordinate the photography, which was configured using a field computer by photo acquisition software to monitor the photography (checking for blur, exposure and distortion) in real time. Images were stored both on an SD card on the camera and relayed to the field computer using a 5 m USB cable through the helicopter door frame.

The flight path was structured into a series of approximately 5 km sections that were orientated around changes in valley geometry. In line with standard practice, images within each section were acquired along quasi-parallel flight lines proving both sufficient side and endlap. Unlike typical UAS surveys, the flight plan was implemented manually by an experienced pilot but was ultimately subject to flying conditions within the complex topography. Mean flight velocity was 80 knots and mean AGL (above ground level) was 350 m resulting in a ground sampling distance of 6.6 cm, similar to the 5 cm achieved by Dietrich (2016). Exposures were configured by software and timed automatically at a frequency of 2 s, resulting in an initial network of 5002 images. The resulting network had a mean side and end-lap of between 50-60% with local variations as high as > 80%, in line with minimum overlap suggestion within the literature (James and Robson, 2012; Fonstad et al., 2013). Due to weather conditions, the flight occurred between 13:00-16:30 (GMT -12) under clear skies. As a result, there were a number of regions prone to heavy contrast from shading in topographically variable regions such as the gorge.

3.3.3.2. The 2015 survey

The 2015 survey comprised a convergent network of images captured using Nikon D750s, one orientated at nadir and one 30^{o} off nadir. These were mounted concurrently

on individual rigid mountings attached to the helicopter skid in a nadir aft-nadir set-up based upon findings detailed later in Chapter 5 (Figure 3.13).

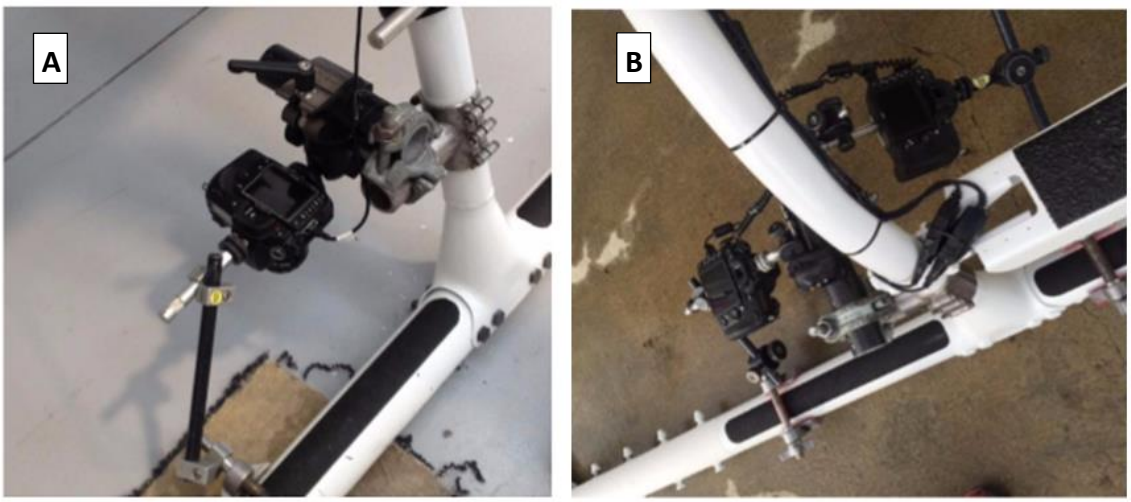

Figure 3.13 Contrast between (A) Singular Nadir orientated camera used in 2014 against (B) Multi-Camera in a Nadir Aft-Nadir set up used in 2015.

Images were again acquired with a fixed 28 mm lens with the focus set at infinity and exposure settings f/6.3, 1/1000s. The Nikon 750 is a full frame camera unlike the DX format of the D90, resulting favourably in a wider field of view. Given the increase in camera sensor resolution (6016 x 4016 pixels), images were not recorded RAW but in a lossless JPEG format giving an average size ~ 15 MB. While this results in a potential loss of fidelity, it was required to manage the computational burden of large file sizes. Following further experimentation documented in Chapter 4, an externally-powered, code-only GPS receiver (a Solmeta Pro 2) was also mounted and connected to each camera. This enabled direct geo-tagging of each image, with WGS84 coordinates and acquisition time written directly to EXIF file of each image. Once more a short test flight was conducted to assess the camera settings and on board GPS recording.

The survey flight was undertaken once more by Glacier Southern Lakes using a Eurocopter AS350 carrying the pilot and an operator for each camera. Once more the flight plan was structured into a series of sections, each approximately 5 km in length, flown at an above ground level (AGL) height of 400 m and at a velocity of 90 knots. The flight plan was orientated around changes in valley geometry, mimicking that from 2014, but accounting for the wider field of view and increased image resolution.

Images were acquired along quasi-parallel flight lines planned to provide both sufficient side and endlap and once more implemented by an experienced pilot, subject to flying conditions within the complex topography. The gorge section consisted again of three transects, however, due to the dual camera rig this resulted in six imaging views, providing sufficient coverage to mitigate some of the problems observed later with the 2014 survey (see Chapter 4). Images were timed automatically using a field computer for each camera and triggered at a frequency of 1 Hz, resulting in a network of >16000 images, with an average endlap and side-lap of 90%. The flight time was scheduled between 11:30-15:30 (GMT -12) to minimize the impact of shadow. As described later Chapter 6 the camera GPS introduced a series of issues that required further processing during subsequent analysis.

The image acquisition strategy for both surveys is summarized in Table 3.4

Table 3.4 Image capture strategies for the 2014 and 2015 Dart surveys.

Protocol	**2014**	**2015**
Camera	Nikon D90	2 x Nikon D750
Camera Settings	Fixed 28 mm lens, f/6.3, 1/1000s	Fixed 28 mm lens, f/6.3, 1/1000s
Image Format	NEF (~40 MB)	JPEG (~12 MB)
Orientation	Nadir	Nadir + 30^{o} aft-Nadir
GPS	No	Yes
Flight Height	350 m	375 m
Flight Speed	80 knots	90 ktnos
Frequency	0.5 Hz	1 Hz
Ground Resolution	6.6 cm	7.2 cm
Images collected	5002	16000
Survey Duration	185 mins	210 mins

Finally, it is worth noting that while the survey strategy used in 2015 was designed deliberately to minimize the incorporation of systematic errors in the self-calibrating bundle adjustment (i.e., improved camera optics, convergent photography, enhanced overlap), these modifications presented significant implications for data processing.

3.4. Structure-from-Motion Processing

3.4.1. Agisoft Workflow

The fundamental theory underlying SfM pipelines has been reviewed in Chapter 2, so here a focus is placed on the implementation of this approach within the software used in this research.

A number of subscription and freeware software have been developed and popularised as tools for SFM photogrammetry. These include Autodesk 123D, Bundler, Mic Mac, Microsoft Photosynth, Pix4D and VisualSfM (Stumpf et al., 2015; Gomez-Guittierez et al., 2014a; 2014b; Kaab et al., 2014; Turner et al., 2012; Woodget et al., 2015). Here, all the subsequent photogrammetric analysis has been implemented using Agisoft Photoscan (APS) Professional Edition, which has been a market leader for over the last five years. One of the key advantages of this proprietary software is the automation of the process, requiring minimal parameterisation of the reconstruction process, while at the same time allowing the user a large degree of intervention if preferred. However, unlike open source codes such as Bundler (Snavely et al., 2006) and PMVS (Furukawa and Ponce, 2010), APS is a comparative black-box, employing its own algorithms and nomenclature that are not documented transparently. Therefore, in the next section, an attempt is made to contextualise the processing steps within APS and relate these to the principles of SfM.

3.4.2. Set-up process

The modelling process in APS is initiated by the creation of a 'project' and the loading of the available imagery and positional data. Manual inspection is undertaken to ensure the photographs used are of acceptable quality (i.e., in terms of image blur and exposure) and do not contain unnecessary large areas of vegetation or wet surfaces which may affect key-point identification and matching (Westoby et al., 2012; Fonstad et al., 2013). The software employs a tool to automatically identify images that have significant blur and fall below a user defined threshold of quality within APS. Positional (including altitude if available) data may be imported automatically if these are written directly within the accompanying EXIF file for each image or imported post hoc through the 'reference pane'. The EXIF data file is also used to provide information on the sensor, lens and focusing.

3.4.3. Alignment phase

The 'alignment' stage within APS refers to the processes of keypoint identification and the initial bundle adjustment to estimate the scene structure (interior and exterior camera orientation and 3D sparse cloud). The keypoint identification enables feature matching, through the examination of neighbourhood pixel intensities (Lowe, 2004) in order to enable the estimation of the fundamental matrix (Snavely, 2006). A non-linear optimization of projection errors is subsequently applied to construct sparse cloud geometry, whereby the maximum likelihood of a 3D point estimate is determined from a set of projection matrices (Rastgar, 2013). Typically, these functions have been carried out by the Scale Invariant Feature Transform (SIFT) algorithm (Lowe, 2004) and Bundler (Snavely et al., 2006), however, APS appears to use a proprietary algorithm.

The execution of an alignment within APS presents a series of available options that relate to both the range of constraints introduced and the effective keypoint resolution and quality. There are four key criteria that may be selected: (1) an 'accuracy setting' that determines the degree of image compression; (2) 'pair-selection' which relates to the degree of positional information; (3) a limit on the number of keypoints that can be detected; and (4) a tiepoint limit places an upper limit on the number of keypoints used and filters these by quality.

(1) The accuracy setting controls the degree of compression that is applied to the raw imagery. In common with many of the parameterization descriptions in APS, a semantic description is used to define the image quality which can be set as high, medium, low and lowest. The high setting supports the use of the native image resolution for the identification of keypoints, while each step below this involves halving the number of image resolution or in effect doubling the pixel size (Agisoft, 2017). A low alignment setting, therefore, results in pixels 16-times larger than in the original image. In practice, this gives a fourfold increase in the mean size of keypoints, as keypoint is a function of not only resolution, but the feature itself.

    The accuracy setting has a major implication for the computational demand of the process and hence execution time. However, reducing the image resolution also affects the density of the sparse cloud, which would have c. only 10% of the

points generated using the low alignment setting (see Figure 3.14 as an example).

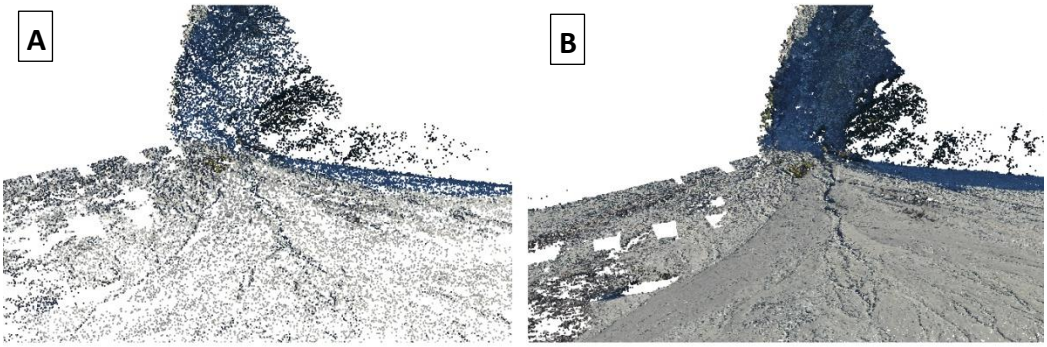

Figure 3.14 Comparison between point densities of SfM reconstructions under (1) Low alignment settings and (2) High alignment settings of the Slipstream debris fan.

For the purposes of this study the High alignment setting has been used throughout. This was chosen to offset the constraints of the survey design, most notably the flying height and speed of the platform, which gives rise to imagery with a comparatively low object space resolution (c. 0.07 m – see Table 3.3). This is significantly lower than many recent SfM studies based on low-altitude UAS surveys and ultimately compromises the horizontal and vertical accuracy of the terrain models that can be derived.

(2) The pair selection setting also has a major impact on the computational time. Two principal settings are available here. For the first of these (pair selection = disabled) images are assessed for correspondence irrespective of any location data provided, resulting in an explosive factorial relationship between image numbers and images pairs tested. The 'reference' setting is used to inform the image matching, so that correspondences are sought between images that lie in similar geographic space based on any positional data added as exterior constraints on the solution. The implications of this are significant as increasingly large image sets result in a non-linear increase in processing times, which both

slow reconstructions and require the acquisition of workstations of high computational power.

For this study, wherever possible, the Reference setting has been used, principally to reduce computational time, enabling the reconstruction of larger datasets involving more images. An additional benefit of providing reference data, is the opportunity to refine the accuracy of the bundle adjustment through these exterior constraints. This results in a smaller parameter space to search and a lower likelihood of incorrectly identifying the unique solution due to the presence of multiple sub-optima in the cost-function.

In addition to the 2D position, information on camera height can be provided or fixed as constant in the reference pane (Agisoft, 2017) and the bundle adjustment can also be refined by providing an estimate of the positional accuracy of the camera data which serves to act as an additional weighting coefficient. A default value of 10 m is selected to match the accuracy of the low-grade code-only GNSS used in most camera georeferencing systems.

(3) The keypoint limit setting determines the upper limit for the number of keypoints identified within an image during the processing stage (Agisoft, 2017). A 'no limit' setting is available, however, the keypoints are filtered by their estimated quality, so the larger the set defined, the lower the reliability becomes. This option was only introduced in recent versions of APS and was not available in the versions used to reconstruct the 2014 models and the laboratory experiments reported in Chapters 4 and 5. Consequently, in the later versions where this was available this limit was set to 40,000 keypoints per image which was the standard value in the earlier versions.

(4) The tiepoint limit determines the upper limit of points per image for the correspondence analysis. This can vary from 0 (unconstrained) to a user defined value, typically of 4000 (Agisoft, 2017). It is imperative to have a sufficient number of matches to enable the construction of the depth maps latterly when building the dense cloud. This study will, therefore, use a value of 0, allowing the

use of all identified matches. Whilst this adds computational cost, there is the opportunity to refine the matches identified at a later stage in the analysis in order to refine the estimated scene structure and remove anomalies using a 'gradual selection' tool.

As an evolving piece of software, APS has since introduced the option to undertake adaptive camera modelling, however, this has not been used at any point during this thesis. Typically, the resulting output from the alignment phase is a sparse cloud comprised of tiepoints in non-dimensional 3D space. However, under the reference settings, the point cloud is projected within the specified co-ordinate system, however, at a coarse degree of accuracy.

3.4.4. Introduction of Ground Control Point network

Where camera positional data have been introduced (i.e., using the reference setting) the resulting sparse cloud generated by the 'alignment' will automatically be projected into that coordinate system. However, given the accuracy of these camera positions and their weighting in the bundle adjustment, it is likely that the result is only approximately georeferenced and may, furthermore, incorporate more complex non-linear distortions associated with errors in the camera calibration. Given the low-quality of the camera positions typically used (based on code-only low-grade solutions), accurate rectification of the model requires control data of a significantly higher order of accuracy (Clapuyt et al., 2016). This is most typically provided by a network of ground control distributed throughout the study area, measured using a precise survey method such as RTK, PPK or static GNSS.

The task of distributing and measuring control was detailed previously in section 3.3.2 and the following section, therefore focuses on the identification and use of control within the APS pipeline. The process of locating targets can achieved manually and/or automatically depending on the target type used and could also be achieved using third party software such as James and Robson's (2012) SfM_GeoRef software. In this study, given the high variability in image contrast and object-space resolution resulting from the range of flight heights, speeds and exposure used, automating this analysis would not be possible (see Figure 3.15) and all targets here were identified manually.

In this process, the GCP co-ordinates are loaded within the reference plane and a new marker is created for each point. The predicted location of each target can then be used to filter the camera list to identify images that should incorporate each marker. Where the scene structure is well-posed, the correspondence between the control and the actual GCP target should be close and only minimal manual adjustment is required to match the predicted and observed locations. As matches are confirmed, the model is continuously updated to facilitate future matches as they are added. This updating is based on a rigid body (affine) transformation adjusting scale, rotation and translation.

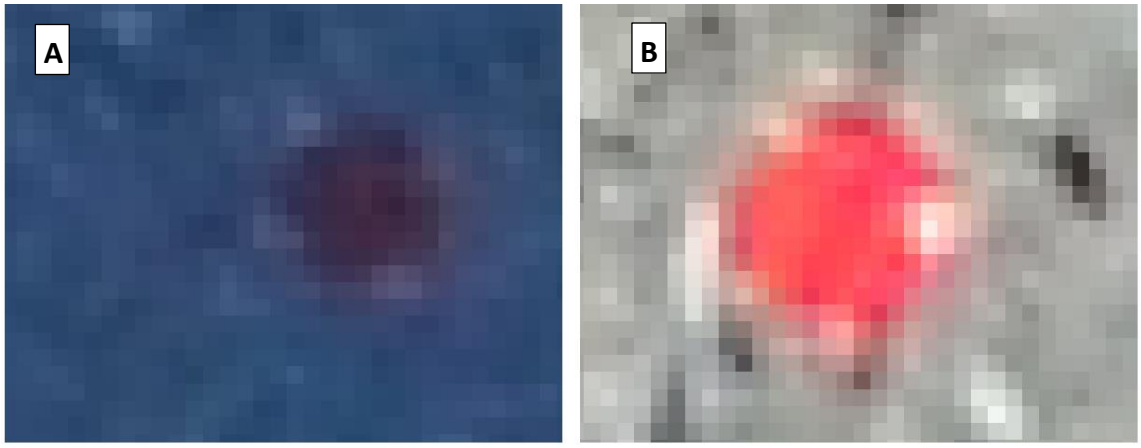

Figure 3.15 Contrasting view of 1 m x 1 m GCPs at differing heights, but an equivalent zoom, (A) Unclear GCP target un-usable, (B) Sufficiently clear GCP target for use.

Ultimately, GCPs need only be identified in two images, however, the accuracy of the solution increases as more images are tied to each constraint (Agisoft, 2017). Once all GCPs have been identified in all available images, the initial scene structure is re-estimated using the newly introduced ground control network with its accurate positional data.

Following this first update, the structure can be further optimized by re-estimating the parameters of the interior camera calibration model. Prior to this step, however, some manual analysis of the structure is advisable, in order to remove anomalies and enhance the reliability of the structure by filtering the structure. This is achieved using a selection tool, which enables thresholds to be applied for key variables and metrics of the solution, including image count, reprojection error (RPE) and reconstruction uncertainty, see Table 3.5 (Agisoft, 2017).

Table 3.5 Major metrics within the APS gradual selection tool for point cloud refinement. Source: Agisoft (2017).

APS metric	Definition
Reprojection Error	The distance between the projection of a 3D point on an image and the original projection of that 3D point detected on the photo. High RPE's tend to indicate poor localization accuracy between correspondences at the point matching step and potential false matches. Hence, the removal of points containing high RPE's can improve the accuracy of optimization.
Reconstruction Uncertainty	The ratio between X/Y and Z error, where high reconstruction uncertainty is typical for small baseline distances between images. This typically introduces noise into the point cloud. Unlike reprojection error, reconstruction uncertainty does not affect the accuracy of optimization. The removal of these points however aids in the visualization of the cloud, particularly useful when checking for surface consistency (see Chapter 6)
Image Count	The number of images used to project the point within the sparse cloud. APS will reconstruct all points visible in at least on two images, however lower numbers are associated with points of poor accuracy and low reliability.

Each tiepoint (i.e., the matched keypoints comprising the sparse cloud) contains information relating to each variable and may be selected for removal based on a user defined threshold. The RPE is often used here, as it relates to systematic distortion resulting from inaccurate camera calibration (James and Robson, 2012). Given the very large number of tiepoints comprising the sparse cloud (unlike standard photogrammetry), the solution is robust to such decimation and refinement. In the case of the 2015 survey described later in Chapter 6, nearly half the sparse cloud was eliminated to refine the estimated structure.

The removal of these points excludes them from the subsequent optimization of the camera calibration which then follows, in order to introduce exterior constraints directly within the bundle adjustment and not just to filter the images matched as described above. This is an important stage, as prior to this, the bundle adjustment is solved purely using inner constraints from the image data and simultaneous estimation of both interior and exterior orientation may result in compensation between these two aspects of the solution. Optimization adjusts the estimated tiepoint coordinates and both the interior and exterior camera parameters, through minimization of the sum of RPE's and reduction in error through reference coordinate misalignment. This is a non-linear transformation that is spatially dependent on the distribution of control points. During this, the user can specify which parameters are free to adjustment, i.e., the focal length,

principal point offset, radial distortion coefficients, tangential distortion coefficients and skew coefficients. This study has focused on the re-estimation of the focal length, principal point, radial distortion (k1, k2 and k3) and skew (b1 and b2) parameters, once more to ensure consistency with earlier editions of the software. The resulting adjustments during the optimization phase are evident in both the interior and exterior pose estimations along with a significant reduction in GCP residuals, producing a more spatially consistent surface.

3.4.5. Pixel-wise densification

The geo-referenced structure may be used as a basis from which to build a dense point cloud, using a pixel matching approach. The majority of studies within the literature refer to this process as Multi-View Stereo or MVS (Westoby et al., 2012; Smith et al., 2015), however, there are wide a variety of densification algorithms. Whilst MVS has become synonymous with this process, APS uses a pairwise Semi Global Matching (SGM) algorithm (Verhoeven et al., 2012). Described first by Hirschmuller (2005), the SGM refines traditional pairwise algorithms by introducing additional points of view and implementing a two-tier penalty system during cost aggregation to distinguish between direction changes and outliers. The resulting summation of strings (points of view) enables the derivation of the minimal cost and disparity map for each pair of images. These disparity maps are then fused (merged and averaged) to produce a final depth map, using a statistical confidence interval to filter overlapping estimates in 3D space, before a Levenburg Marquandt approach is used to minimize geometric error in the cloud (Hirschmuller, 2005; 2008). The SGM approach is computationally advantageous as the structure of algorithms is suitable for parallel processing allowing the use of multiple CPU's and GPU as implemented within APS.

Within APS, the densification process utilises the sparse cloud to identify suitable imaging pairs and provide an initial seeding based upon the number of matches between pairs. These seedings are then used to establish the order in which the pairs are incorporated into pairwise SGM algorithm. The camera calibration (both interior and exterior) remained constant throughout the process and are used to define 16 epipolar lines through each image pair for the process of cost aggregation and disparity mapping. Within APS there are once more two settings, one relating to pixel size and the other the degree of surface filtering. The medium density setting has been applied throughout

this study as it is capable of providing point cloud densities sufficient to enable the creation of sub meter resolution grids (Figure 3.16).

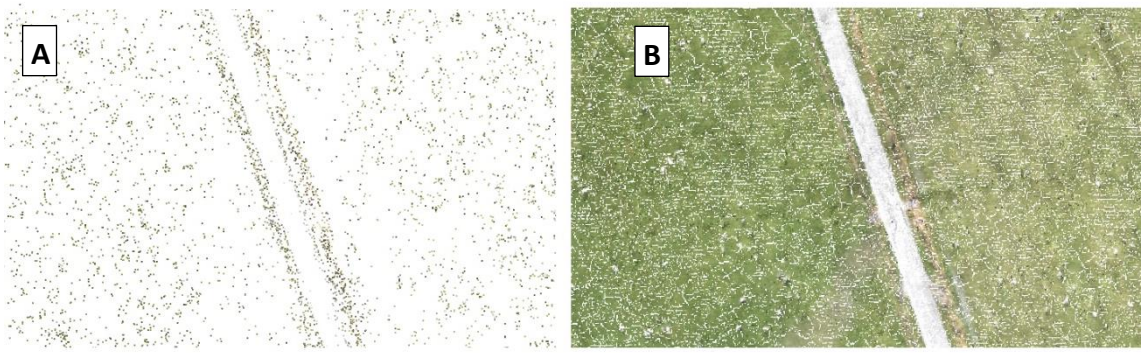

Figure 3.16 Comparison between (A) sparse cloud consisting of tiepoints and (B) dense cloud performed under medium settings of a 100 m stretch of road.

Alongside this, an aggressive filter has been selected which influences the degree of surface smoothing in the final derived point cloud. This approach results in some loss of surface detail, as this is followed by a further decimation and conversion to a raster DEM (see section 3.6), however, the aggressive technique provides a means to reduce anomalies that may distort later terrain products. Whilst APS remains a black box piece of software, it may be assumed that the filtering level applies to the two tier penalty system employed during cost aggregation, either altering the threshold of pixel boundaries, or modifying the magnitude of penalty. The final dense cloud produced has broadly the same geometry as the sparse cloud, however, tests have shown this to be prone to less surface noise as a result of the smoothing semi-global approach to depth map fusion. The final point density is a function of both the dense setting chosen and the pixel resolution of the imagery.

3.4.6. APS Summary

The following Table (3.6) provides a summary of settings used within APS and how these relate to key steps of the SfM pipeline. Depending on the available computational power, other studies may be required to use higher or lower alignment settings.

Finally, it should be noted that the approach described here and used in the subsequent chapters has been tailored to facilitate the management of the very large datasets used

in this study (i.e., 16,000 x 24 MP images), while relying on available desktop computing. It is also important to recognise that the geometry of the reconstruction is determined during the sparse alignment and optimization phase. As such the majority of focus of any APS based workflow must be aimed at optimizing the quality of the sparse cloud.

Table 3.6 A summary of the key settings selected within APS during the majority of reconstructions undertaken within this thesis.

Phase	Alignment Value	Alignment Type	Gradual Selection	Optimization	Dense Alignment
Setting used within APS	High, 40,000 keypoint limit, 0 tiepoint limit	Reference	RPE <1, Reconstruction uncertainty <10	Fx, Fy, Cx, Cy, K1, K2, K3, B1, B2	Medium, Aggressive
Relation to Pipeline	The resolution of imagery and resultant keypoint size. High alignment setting allows the full image resolution to be exploited. In comparison to lower alignment settings, this enables the identification of additional keypoints and prevents their fidelity within space being compromised with downscaling. This is of importance to imagery collected at both increased heights and flight speeds, where pixel fidelity may be reduced.	This identifies predetermined imaging pairs, reducing computational time and providing additional exterior constraints for the bundle adjustment. In comparison to the disabled setting, the Reference setting allows the use of additional imagery or increased alignment value settings, vital to larger survey areas of more complex geometries.	The removal of spurious tiepoints that may be incorporated further into bundle adjustment and delineation of point cloud surface to enable manual quality control. Minimising RPE is likely to improve the results of the optimization phase, whilst noisy surfaces associated with convergent imagery with short baselines may be cleaned up to detect regions of discontinuity.	The final step of bundle adjustment, incorporating ground control to improve calibration estimates through improved exterior constraints. The settings chosen reflect those available in earlier editions of APS that were used to perform reconstructions within Chapter 5.	The implementation of a pairwise SGM that results in dense cloud through pixel-wise matching. The setting chosen has no impact on reconstruction geometry as the calibration remains fixed. Instead the setting chosen relates to the desired point density and likely computational limitations. The aggressive settings is used within this study as the focus of this study is on broad systematic elevation changes as opposed to specific feature change.

3.5. Additional Datasets for Model Design and Validation

For both the 2014 and 2015 surveys, additional data were obtained to provide spatially continuous control; for the distribution of these data see Figure 3.17. The protocol used to obtain these datasets is detailed in the subsequent sections.

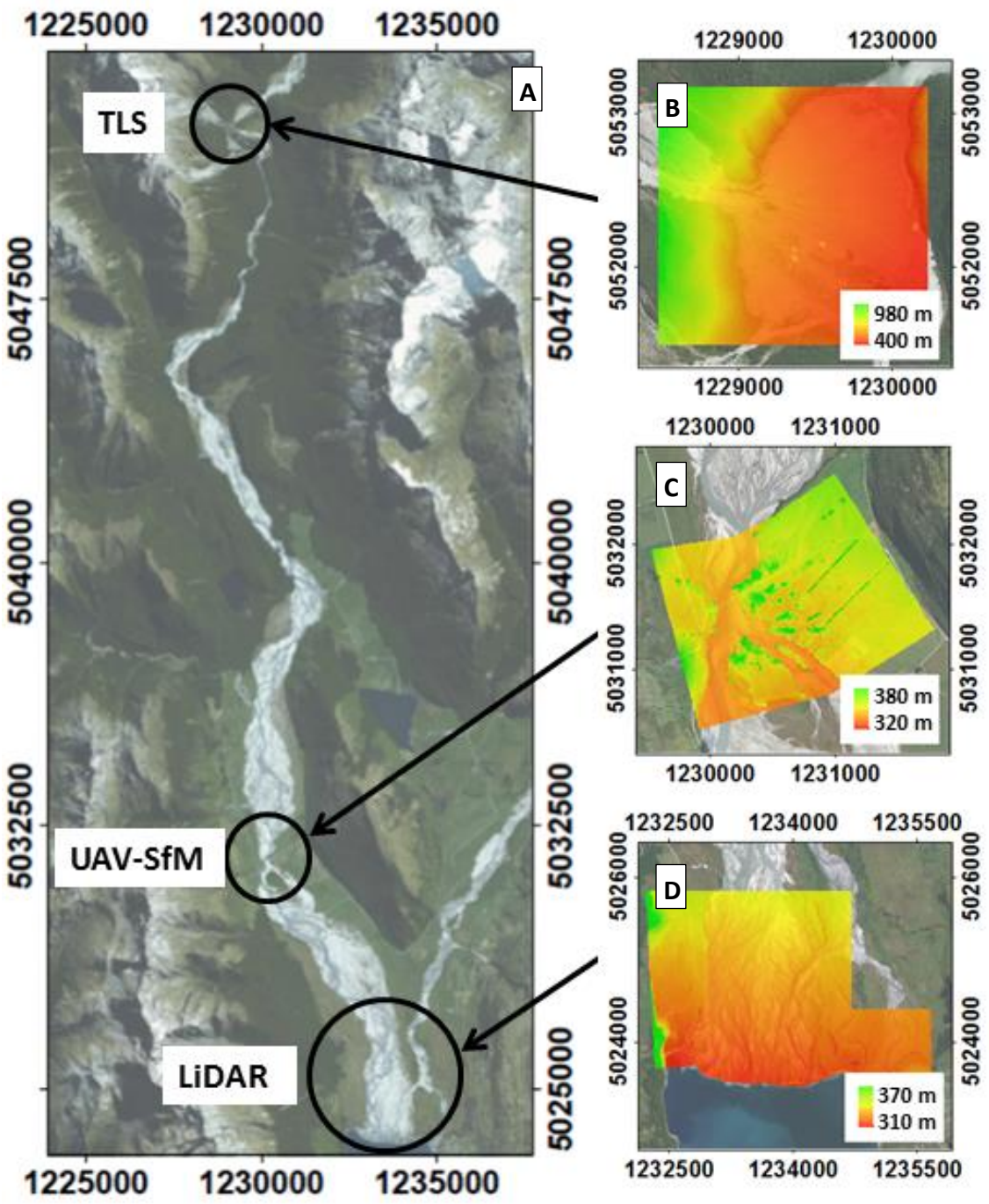

Figure 3.17 (A) Location and DEMs produced for each control dataset used within this study. (B) TLS was collected in 2014 as part of the original field campaign, (C) the UAS-SfM by the University of Otago in 2015 (Sirguey, 2015) and (D) The LiDAR was obtained from a 2011 survey commissioned by the ORC (NZAM, 2011).

3.5.1. Laser Scanning

Independent assessment of the 2014 survey was supported by comparison of the derived DEMs against datasets obtained using airborne and terrestrial laser scanning. Laser scan data offer the opportunity to gain an insight into the spatial distribution of errors incorporated in SfM models which are unlikely to be well described by a small set of independent check point measurements. A number of previous SfM studies previously have used spatially explicit check data in this way, most commonly based on airborne laser scanning (e.g., Fonstad et al., 2013) and increasingly with TLS (Westoby et al., 2012). It should be recognized that such datasets, while of comparable spatial density, are not without errors themselves and should not, therefore, be considered 'ground truth'. Both airborne and terrestrial laser scanning are prone to registration (whether calculated on-the-fly with an INS, or inversely by resection from targets) and ranging errors and more complex problems such as multi-pathing and multiple reflections (see Brasington et al., 2012).

However, while not 'true' representations of the surface, such dense data models are likely to be of similar if not greater reliability than medium altitude SfM and the redundancy provided by the density and coverage of measurement provides a good basis for model assessment. Datasets derived by ALS and TLS are also more familiar to many users within the geomorphological community and so provide recognizable data with well understood uncertainties. In this study, a combination of both historic airborne laser scanning (hereafter LiDAR) and specifically acquired TLS have been used.

3.5.1.1. LiDAR

Airborne laser scanning typically incorporate a near infrared laser scanner with positional and attitude data provided by a GNSS-augmented inertial navigation system (see Shan and Toth, 2009 for more background). Data are most commonly acquired from fixed wing platforms flying at between 1000-3000 m flying heights. In this study, a LiDAR survey of the lower 2-3 km of the Dart-Rees delta including the township of Glenorchy was made available by the Otago Regional Council (ORC). This survey was captured for the Queenstown Lakes District Council by NZ Aerial Mapping and Aerial Surveys Limited (NZAM, 2011). The data were acquired between the 30/09/11 – 03/10/11 using an Optech ALTM 3100 EA LiDAR system flying at 1,200 m AGL with field

of view of 42 degrees. Ground referencing was achieved using geodetic marks at Wanaka and Kingston airfields. The data were post-processed using Applanix POSPac software and coordinates defined using NZGD2000. Both digital surface and bare earth digital elevation models were made available at 1 m ground-spacing in raster format. Only limited information on model assessment was provided, based on height difference statistics between TIN and LiDAR ground points at an unspecified number of checkpoints. This analysis suggested typical standard deviation statistics of 0.026 m at Glenorchy, ranging up to 0.051 m at Wanaka.

This data set is used to identify broad systematic trends in the derived SfM models, albeit in a very marginal region of the reconstruction. Moreover, given that 3-4 years had elapsed between the LiDAR and 2014/15 SfM surveys, there is likely to have been significant disturbance in the intersecting areas of the active channel belts of both the Dart and Rees rivers, so that the most effective comparison is restricted to stable regions of Glenorchy itself.

3.5.1.2. Terrestrial Laser Scanning

In 2014, as part of the same field campaign used to acquire ground control and imagery for the SfM survey, a Reigl VZ-1000 long range TLS scanner was used to generate a point cloud of the debris fan and landslide at Slipstream (Figure 3.18).

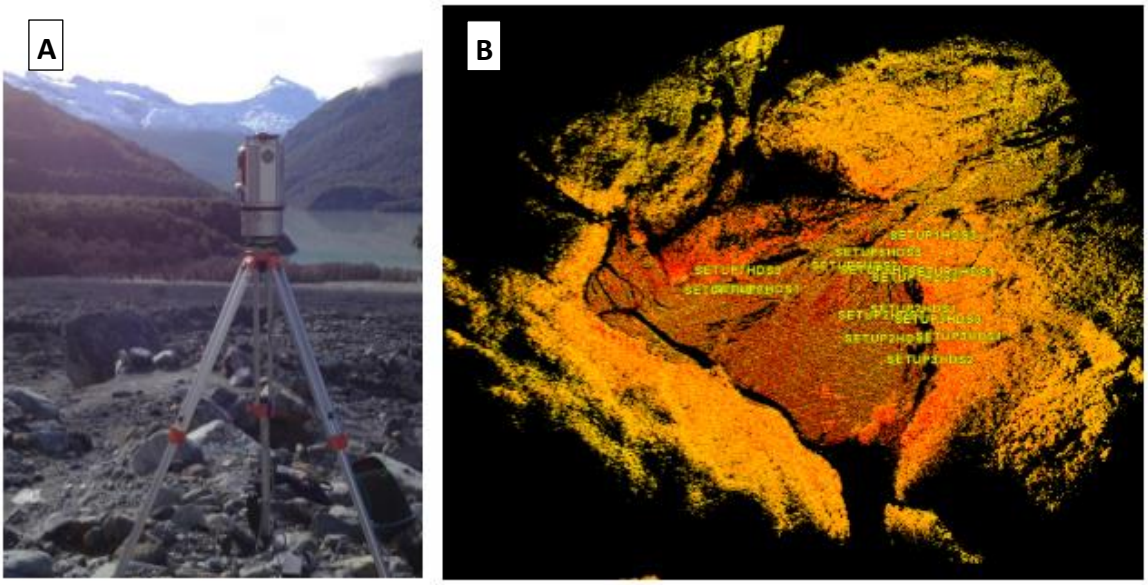

Figure 3.18 (A) Reigl VZ-1000 long range TLS scanner and (B) Unified point cloud with target locations within Leica Cyclone.

The Scans were undertaken over a two day period of dry conditions on the 10/05/14 and 11/05/14, three days ahead of the aerial survey on the 14/05/14. Scans were obtained from seven positions across the debris fan, and geo-referenced by three tripod mounted targets with positional data acquired using RTK-GNSS. A total of 21 control points were used. Post processing was undertaken using a combination of RiSCAN PRO and Leica Cyclone software. The seven scan data sets were co-registered into a single point cloud through using a global 3D similarity transform based on the tripod targets. The mean absolute registration error was found to be -0.103 m, somewhat higher than would be expected for a relatively well-constrained TLS survey. This reflects problems with tripod stability for both the registration targets and in particular the TLS itself, which were prone to sinking in the very soft, wet sediment comprising the debris fan following recent activity.

3.5.2. Hillocks UAS-SfM

A further independent survey was acquired to support assessment of the SfM modelling undertaken here. This was based on low-altitude UAS-SfM survey undertaken by project partners from the University of Otago, New Zealand, incorporating a 2.8 km^2 area of the Hillocks at the Dart River crossing (Sirguey, 2015). The survey was acquired on the 19/05/15 with a Trimble UX5 remotely piloted aerial system operating at a flight height of 122 m AGL and imaging with a Sony NEX-5R camera and 15 mm f/4.5 lens, giving a ground sampling distance of ~ 0.04 m.

Figure 3.19 Trimble UX5 fixed wing UAS with range of 60 km and max flight time of 50 minutes.

The survey comprised two flights and a total of 39 flight lines, obtaining 2340 overlapping images of 73% side lap and 80% end lap. Imagery was referenced with 32 60 cm x 60 cm targets, painted to contrast the surrounding area for ease of location. In line with the procedure in section 3.2, targets were measured with a GNSS rover mounted to a survey pole collecting observations in RTK from a base station post-processed using MAVL and LEXA 0 order benchmarks. The imagery was processed in Trimble Business Centre (TBC) v3.40 generating a point cloud of 2.12×10^8 points, with a mean point density of 79 pts m^{-2}. Subsequently, the point cloud was interpolated to produce DEMs at a resolution of 0.15 m. Estimates of model accuracy are provided by 27 GCPs with a RMSE of 0.028 m with no dimensional bias. These points are part of the bundle adjustment and so act to constrain the resultant surface, thus a total of 15 Independent Control Points (ICPs) were surveyed, with a 3D RMSE of 0.094 m.

3.6. Deriving Topographic Data Products

The datasets obtained by the TLS and SfM surveys (2014 and2015) produce dense point clouds that may comprise 10^6–10^7 individual observations. Geomorphological analysis of these data is greatly facilitated by their conversion into raster terrain models, which can be simply and quickly manipulated and visualized within a GIS (Vericat et al., 2014). The following section describes the approach used here to convert the point cloud data into bare earth DEM models.

3.6.1. Point Cloud Decimation

Each point cloud was decimated and regularized using the open-source Topographic Point Cloud Analysis Toolkit or ToPCAT (Brasington et al., 2012), which is now incorporated as part of the GCD software. As described in Chapter 2, ToPCAT samples the point cloud using either a 2D grid or 3D voxel and determines the local mean, minimum and maximum elevation in each cell, along with a statistical summary of the local cloud. This initial filtering of the cloud is then used to define a global tessellation of the point cloud at the resolution of the sampling grid, which in turn is used to detrend the local statistical summary. While developed principally for use with TLS data, ToPCAT has been implemented successfully on previous SfM studies across a range of landform types; analysing coastal cliffs, a moraine dam, and glacially sculpted bedrock ridge

(Westoby et al., 2012), fluvial morphology (Javernick et al., 2004; Marteau et al., 2017) and Badlands (Smith and Vericat, 2015).

3.6.2. DEM Generation

All DEMs were subsequently interpolated using ArcGIS 10.4 following the same workflow and implemented within the Model Builder tool. This was achieved by interpolating the local minimum elevation of the reduced resolution point cloud derived from ToPCAT to generate a TIN and then resampling this vector data model to a regular grid and using the default linear interpolation. The resulting raster was then clipped to a consistent area of interest, taking care to ensure that all grids met standards of orthogonality and concurrency in order to support later analysis.

3.6.3. DEMs-of-Difference

In line with the approach pioneered by Lane et al. (1994), repeat DEMs can be simply subtracted to provide estimates of geomorphological change in the form of a DoDs. This can be implemented using standard raster logic tools or supported by plug-in software such as the GCD software (Wheaton et al., 2010). As described in Chapter 2, changes observed between datasets are likely to incorporate both real geomorphological changes and also spurious differences that reflect errors in both terrain datasets. In much of the change detection work that follows, a simple approach to modelling was employed based on the specification of a Minimum Level of Detection, minLOD (Brasington et al., 2003; Lane et al., 2003; Wheaton et al., 2010). This is derived from an error analysis of each surface, which is often classified according to surface cover (e.g., wet, dry, vegetated) and the reliability of change detection determined using a minLOD derived from the theory of errors described in Chapter 2 and implemented using the GCD plugin (Wheaton et al., 2010).

3.7. Error Metrics

Throughout this thesis a number of error metrics are used to describe uncertainties in the data products and support DEM comparisons. The range of metrics used is, in part, an attempt to address the inconsistencies in error reporting within the literature and in part, to reflect different characteristics of the error structure that the metrics capture.

A summary of the nomenclature and definition of terms is given presented below and used throughout the following chapters.

3.7.1. Mean Error

Mean Error (ME) is an indicator of survey bias and can be derived from control point residuals and also from the set of errors determined from a DEM to DEM comparison (i.e. a synthetic DoD comparing to models of the same surface). It is defined as:

$$ME = \frac{1}{n}\sum E \qquad [3.1]$$

Where $\sum E$ is the sum of residuals and n is the number of observations. The dimensional nature of the equation results in the indication of an overall trend but does not, therefore, provide information on the average displacement.

3.7.2. Mean Absolute Error

Mean Absolute Error (MAE) an indicator of the average magnitude of error without considering their direction:

$$MAE = \frac{1}{n}\sum_{j=1}^{n}|y_j - \hat{y}_j| \qquad [3.2]$$

Where, y_j is the ith observed (check) elevation and $\hat{y}_j$ is the *i*th predicted (modelled) elevation. This metric is useful for providing an estimate of accuracy when errors are likely to be both positive and negative. A distinct disadvantage, however, is the production of an absolute value, which creates difficulties in a number of mathematical equations, such as those used to parameterise models (Chai and Draxler, 2014).

3.7.3. Root Mean Square Error

Root Mean Square Error (RMSE) is a measure of the average magnitude of error that is more sensitive to large outliers:

$$RMSE = \sqrt{\frac{1}{n}\sum_{j=1}^{n}(y_j - \hat{y}_j)^2} \qquad [3.3]$$

The RMSE is typically used to evaluate model performance within the geosciences, howeve, the assumption is that errors are normally distributed (Chai and Draxler, 2014). The term incorporates both systematic bias (see mean error) and random variation around the average error.

3.7.4. Standard Deviation

Standard Deviation of Error (SDE) is an indicator of spread within the modelled residuals, separate from any systematic bias that might exist and affect terms such as the RMSE. This is defined as:

$$SDE = \sqrt{\frac{1}{n}\sum\left[\left(y_j - \hat{y}_j\right) - MAE\right]^2} \quad [3.4]$$

The Standard Deviation (SD) is used to assign a value to surface variance at points (Chapter 6 and 7) is;

$$SD = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(x_i - \mu)^2} \quad [3.5]$$

In this instance, the SD represents the spread of data about a mean value, which is assumed to represent the true surface.

3.7.5. Max Error

The Maximum Error (MaxE) is presented in relation to both GCP and ICP residuals within this thesis and is a measure of the error of most significant magnitude. This may be used to represent the potential for error within a survey and is most likely associated with errors resulting from systematic bias inherent within the data as opposed to local noise.

3.7.6. Cross Validation Statistics

When calculating errors based on a sparse network of control points, two approaches are commonly adopted to separate errors associated with controls used in the modelling process (and are therefore biased) and those that are independent. The most common approach to this is the Leave-p-out cross validation, in which p observations are excluded from the training data and used to assess the model fit retrospectively. For example, for a SfM model with a total of 100 GCPs, this might involve randomly sampling

20 GCPs to define a broad test of the model, then using the remaining 80 to fit and optimization the bundle adjustment. Errors would then be assessed using the retained 20 GCPs to provide an independent check on the model.

While Leave-p-out cross validation provides a strong independent assessment of model performance, it suffers from two key problems. First, the p observation set are fixed and represent a static test, thus they may not be representative (spatially or statistically) of the wider model. Second, by excluding p observations, the model itself does not make optimal use of all available data.

An alternative cross-validation strategy is Leave-one-out cross-validation (LOOCV). This approach involves removing one observation as an independent check observation and computing the solution. This observation is then returned to the training set and another observation excluded and the model refitted. This process continues, cycling each observation in turn until a set of model errors can be defined from each observation in turn. This approach offers the benefit that the final set of check observations include all available data (spatially and statistically) and thus represent the most effective sampling density. Furthermore, the fitted model tested makes optimal use of all available data. The downside of this strategy is that it requires extensively remodelling. For example, taking the case described above with 100 GCPs, LOOCV would require re-estimating the bundle adjustment 100 times to generate an independent solution for each GCP as a check point.

3.7.7. Summary of error metrics

The choice of multiple metrics enables an assessment of their respective ability to represent the potential magnitude and spatial distribution of errors. Both the 2014 and 2015 survey experience regions of limited ground control that are inconsistent between years; a trait that is symptomatic of catchment scale SfM surveys. When considering the two most commonly used metrics (MAE and RMSE) the latter is more sensitive to outliers, typical of systematic errors within SfM. However, both the RMSE and MAE are averaged across all residuals, including those within the boundaries of ground control, resulting in a diminished value for error. In contrast, the MaxE is a unique value, providing a more realistic value for the true potential for error.

4. Chapter 4: Topographic reconstruction of the Dart River, 2014: Evaluating the capabilities of the Structure-from-Motion pipeline to answer catchment wide morphological hypotheses.

Abstract

The structure-from-motion (SfM) pipeline has the potential to capture topographic data over significant spatial scales to address catchment wide research questions. The application of this emerging technology at this scale urgently requires robust testing and validation. Results are presented here that evaluate the 3D reconstruction of the Dart River at a range of scales based on data acquired in 2014. This survey comprised nadir orientated imagery from a digital SLR (DX format) camera, captured from a helicopter at an above ground level of 350 m and flight speed of 85 knots. Residuals for Ground Control Points (GCP) have a Mean Absolute Error (MAE) of 0.061 m, up to a maximum of 0.622 m, while Independent Control Points (ICP) distributed across the reach were found to have a MAE 0.441 m, with a maximum value of 2.668 m. A comparison of the model of the delta against an existing LiDAR survey from 2011, reveal that 71 % of stable areas deviate by less than 0.5 m of the LiDAR data within the area of the GCP network however, deviations of 5 m are observed beyond this. A detailed study of a well-controlled site, a debris fan at Slip Stream, was enabled by acquiring a supplementary model of this area using long range Terrestrial Laser Scanning (TLS). This survey was used to provide dense spatially distributed data for more robust model validation. In this analysis, 89% of area within ground control was found to have vertical residuals of <0.5 m (MAE 0.303 m). Additional control points at the perimeter of the area of interest were incorporated and found to exhibit errors of between 0.8 m to 2.6 m, while DEM-to-DEM comparisons revealed significant distributed patterns of difference, indicative of systematic error.

Strategies to minimize errors are evaluated by introducing a series of refinements at various stages of the SfM photogrammetry workflow. The inclusion of strategically-placed additional ground control was found to create a step change in quality of results, while incorporating inferred camera position information was found to reduce the computational overhead, enabling large single block reconstructions. The block bundle

adjustment was shown to be robust to random errors of up to 10 m in the *a priori* estimated camera positions, which represent typical positioning data from code-only OEM receivers on commercial cameras and UASs. Results from both the catchment wide survey and debris fan, highlight the importance of effective photogrammetric camera networks to provide robust solutions to the bundle adjustment with limited exterior constraints (ground control).

4.1. Chapter Objectives

This chapter reports on research which aims to evaluate the potential to derive broad-area, high resolution (sub-metre) digital elevation models (DEMs) with minimal systematic bias, suitable for supporting catchment-scale geomorphological research and specifically geomorphological change detection modelling. A driving goal is the requirement to generate reproducible DEMs that facilitate detection of the typical topographic changes associated with channel adjustment following the introduction of a significant sediment pulse (Lisle et al., 2001; Cui and Parker, 2003; James, 2010). An important consideration in this context is the presence of systematic, spatially-correlated DEM errors that conventional quality statistics (e.g., the SDE, MAE) derived from sparse networks of ground-control and independent check points often fail to reveal. DEM differencing techniques are particularly sensitive to such systematic errors and these can lead to widespread spurious estimates of surface change, especially when the structure of errors differs between the two survey datasets (see section 2.6). Structure-from-Motion derived DEMs have previously been shown to be highly sensitive to systematic bias, arising from compensation between the exterior and interior orientation parameters. Unlike traditional photogrammetric methods, in SfM the inner and exterior geometry are solved simultaneously in a self-calibrating bundle adjustment most commonly without any exterior constraints on the solution (i.e., ground-control or measurements of camera position and pose). These effects are further exacerbated by the frequent reliance on consumer-grade cameras that have unknown and potentially unstable interior geometry.

A common strategy to offset these effects is the use of extensive ground control to optimize the results of a prior bundle adjustment with a sparse network of exterior constraints (James and Robson, 2014). While this approach is relatively straightforward for surveys that extend only over small areas (< 1-2 km^2), deploying ground control in remote, inaccessible terrain presents a significant logistical challenge. A critical goal, therefore, is to develop strategies to first quantify the presence of systematic errors, and then consider methods to mitigate these effects to upscale the SfM workflow to the catchment scale. In this chapter, this broad goal is addressed through the following related objectives:

Objective 1: Develop a SfM workflow to generate a 3D reconstruction of a 35 km reach of the Dart River using nadir orientated imagery acquired in 2014.

Objective 2: Quantify the magnitude and structure of uncertainty in the resulting DEM.

Objective 3: Conduct an experimental analysis of uncertainties generated within the APS SfM pipeline using dense independent data from a localized area of interest, the debris fan at Slipstream, to provide benchmark topographic data.

4.2. Reconstructing the 35 km reach

4.2.1. Image Acquisition

The image acquisition strategy for the 2014 survey is described in detail in Chapter 3. Briefly, images were obtained with a nadir orientated Nikon D90 DX SLR camera attached on a rigid mounting, attached to the helicopter skid. Images were acquired with a fixed 28 mm lens with the focus set at infinity and exposure settings f/6.3, 1/1000s with images recorded in RAW, NEF format. The flight plan was structured into a series of sections, each approximately 5 km in length that were orientated around changes in valley geometry (Figure 4.1). Within each section, images were acquired along quasi-parallel flight lines proving both side and endlap. The flight plan was implemented manually by an experienced pilot but was ultimately subject to flying conditions within the complex topography. An intervalometer was used to trigger exposures at a two second frequency, resulting in a network of 5002 images at an above ground level (AGL) of 350 m and an average speed of 85 knots. The resulting image network had similar side and end-lap of between 50-60% but locally as high as > 80%. Due to weather conditions, the flight was scheduled in the afternoon under clear skies. The set of images, therefore, incorporates some significant local shadowing and high contrast associated with variations in the surface albedo (reflective gravel surfaces vs. dense beech forest cover).

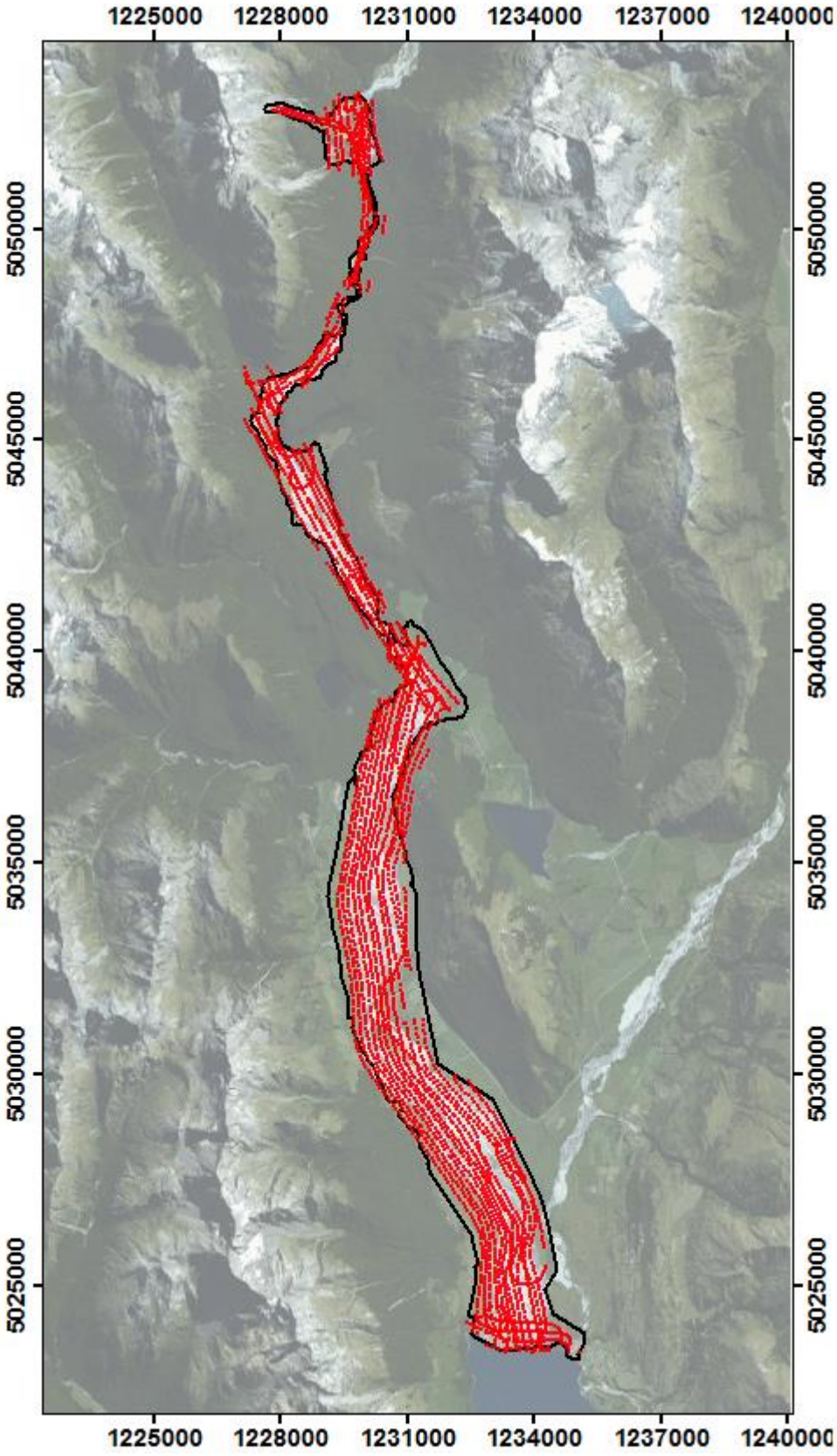

Figure 4.1 Image locations captured for the 2014 Dart SfM survey. The number of parallel transects varies according to valley width, ranging from > 10 per channel width in the delta to just 3 across the upstream gorge. *Projection: NZTM 2000*

4.2.2. Development of a 3D Reconstruction Workflow

The generalised steps taken within APS are presented in section 3.4; alterations to this process during the 2014 campaign are presented subsequently. Reconstruction of the 3D structure, comprising the interior and exterior camera orientation and 3D tiepoints, of the whole study reach presents significant data processing and modelling challenges. These arise from a combination of the environmental and logistical constraints of surveying wide areas of a mountainous, wilderness environment. These combined effects give rise to a dataset that incorporates key compromises that are described in Table 4.1.

Table 4.1 Environmental and logistic constraints on the survey strategy and their associated photogrammetric consequences.

Environmental/Logistical Challenge	Effect on Data Acquisition	Photogrammetric compromise
Large survey extent	Low density of ground control; significant human resources and time required to deploy ground control and acquire validation data; long flight duration (3-4 hours).	Sparse exterior network of exterior constraints; only localized distributed data for model validation; variable illumination due to changing solar angle.
Labile fluvial environment	Ground control restricted to channel margins in areas unlikely to be disturbed during the survey campaign.	Irregular and sub-optimal placement of exterior constraints.
Complex topography	Complex flight planning; variable illumination and significant shadowing.	Corridor photogrammetric network with limited image convergence; variable side and endlap; poor image quality limiting the accurate identification of keypoints and ties.
Contrasting albedo	Strong variation in albedo of surfaces (highly reflective schistose gravels and dark forest canopies).	Combination of over and under exposure limiting the accurate identification of keypoints and ties.
Wetted areas	Complex distribution of wetted areas (braided network).	Inconsistent scene structure between sampled images due to motion generating the potential spurious tiepoints; difficult to mask individual images.
Financial constraints	Limited flight time; use of existing camera with low resolution DX format and no on-board GPS.	Compromise between survey extent and flying speed results in limited side and endlap; limited field of view, further restricting image overlap; low spatial resolution imagery limiting finite size of keypoints and quality of matches.
Computational resources	Photogrammetric modelling on desktop windows workstation (Dual Core I7, 16 GB RAM).	Insufficient processing power and memory to build a single reach-scale block in the absence of camera position information.

To develop a 3D reconstruction of the entire reach, a heuristic methodology was developed that involved a series of successive refinements to meet these challenges and

optimise the results. This series of steps is documented in Figure 4.2 which shows four stages in model construction.

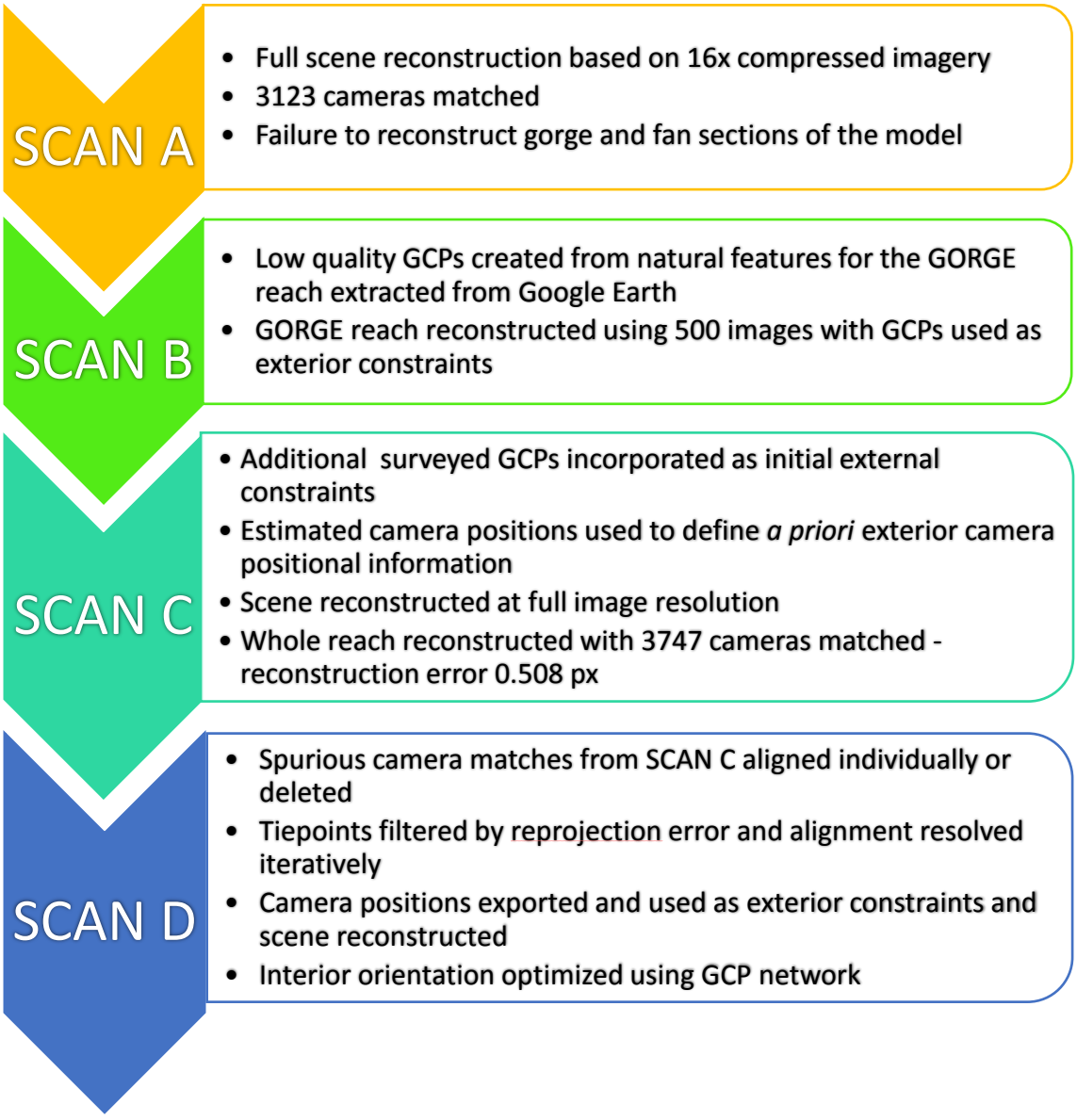

Figure 4.2 Heuristic workflow developed to solve 3D scene geometry for the full 35 km reach. The workflow builds iteratively, incorporating surveyed (GCP) and posterior modelled (camera locations) exterior constraints

4.2.2.1. *SCAN A: Reduced Resolution Model of the Entire 35 km Dart River*

The raw set of photographs comprised 5002 images in RAW format and so constitutes a major computational task to derive the entire structure in a single block bundle adjustment. This approach is, however, desirable as splitting the area of interest into smaller blocks, or chunks as they are referred to in APS, risks incorporating systematic errors particularly at the margins of blocks that then need to be merged to produce a whole, consistent scene.

To address this, the initial set of images was reduced to 4009 by removing redundant images dominated by significant areas of water and dense vegetation which were likely to give rise to spurious tiepoints. A reconstruction was then processed without any exterior constraints (i.e., without information on camera positions or ground control) using the 'low' alignment setting in APS. This approach reduces the computational overhead by resampling the imagery to a quarter of their native spatial resolution, thus in this case, reducing the typical object space resolution from 0.07 to 0.28 m. While it would be desirable to construct the model at the original image resolution, tests on the available workstation, a Dell Optiplex Dual Core I7 with 16 GB RAM and 1GB NVIDIA graphics, identified that this level of compression was necessary to execute a bundle adjustment that encompassed the whole 35 km reach. This initial reconstruction, generated using reduced resolution imagery and with inner constraints only, resulted in an arbitrarily scaled and orientated, but incomplete model consisting of 767,225 tiepoints from 3123 aligned cameras. The entire upstream sections of the model incorporating both the Gorge and Fan reaches were not solved successfully and omitted.

4.2.2.2. *SCAN B: A Local Model of the Gorge Reach*

The failure to solve the entire scene in one block is partly due to the limited information in the set of images, due specifically to the restricted overlap, the low resampled image resolution and, in particular, the dominant effects of shadowing through the Gorge reach which limits the effective identification of keypoints. This is reflected in the sparse cloud in which imagery within the Gorge yields on average 1512 tiepoints per photo, as compared to 3986 per photo for the rest of the catchment. A partial solution to these problems can be found by incorporating additional external constraints as part of the bundle adjustment. These can take the form of either/both interior and exterior orientation of the cameras and/or ground control data. For the Gorge reach there is,

however, limited external information available as: (i) the placement of GCPs was constrained by the limited number of stable sites (gravel bars and small floodplain pockets) that could be seen from above and were free from tree cover; (ii) the images were acquired without GPS positioning so there is no available exterior information on the camera network; (iii) the camera was uncalibrated.

In an attempt to address this problem, a set of natural features (boulders, natural corners) were identified within the Gorge reach (Figure 4.3).

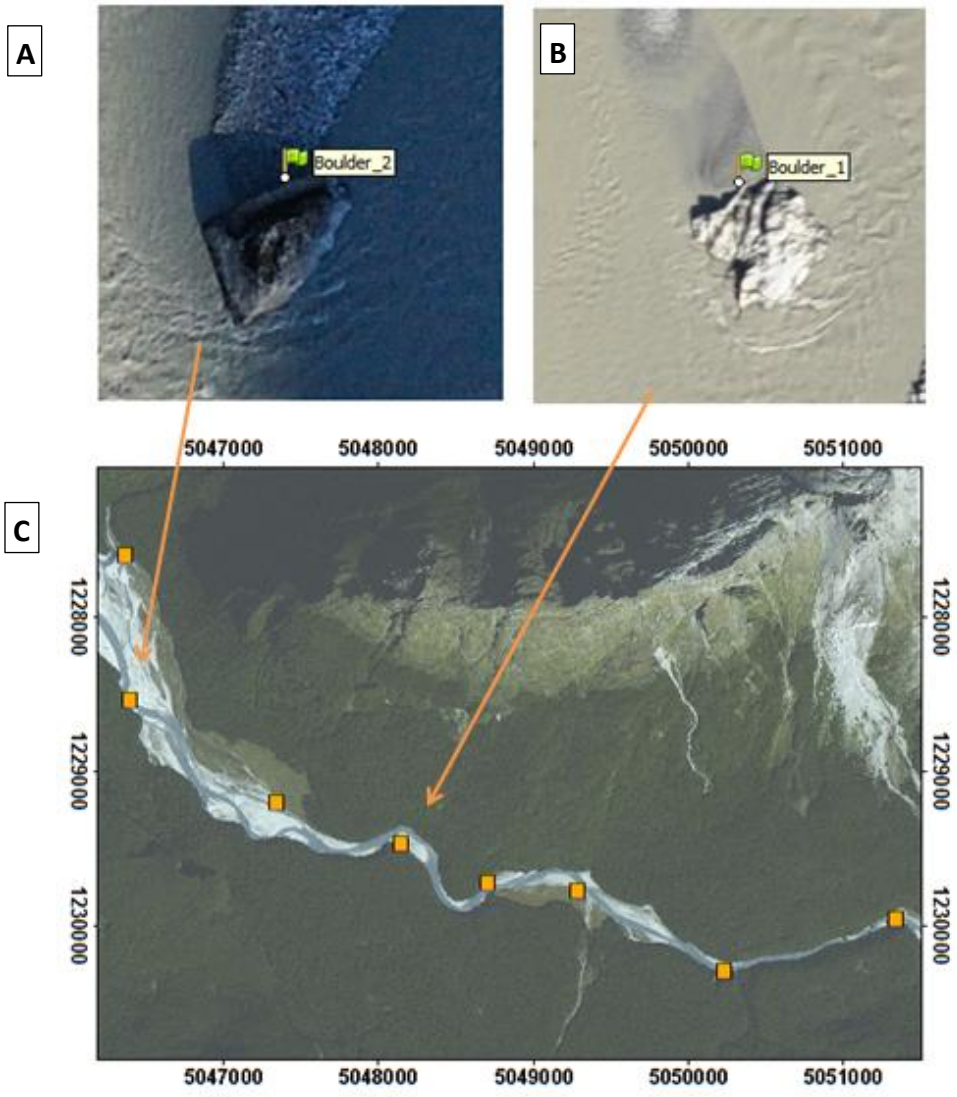

Figure 4.3 Map of the Gorge reach showing the location of placed GCPs and two corresponding natural features identified initially in Google Earth (A and B) and subsequently within the aerial imagery (C) to provide additional exterior control. *Projection: NZTM 2000*

Features were identified using imagery and 3D coordinates obtained from a high resolution satellite image (acquired on the 11th June 2014) available from Google Earth and provided by Digital Globe Inc. This approach was used to create 8 low quality control

points that were incorporated into a new block of the Gorge reach only (n = 500 images) as exterior constraints and set with a low 3D positional reliability of 10 m. The 8 control points were observed in an average of 10 image sightings per marker. In contrast, the original GCP distribution provided 5 GCPs with an average of each marker appearing in just 5.4 images. The distribution of natural features and placed GCPs within the Gorge reach is shown in Figure 4.3. The resulting reconstruction successfully matched 497 cameras with 500,000 tiepoints, with a 3D RMSE based on the Google Earth GCPs of 0.02 m.

4.2.2.3. *SCAN C: A first Block Bundle Adjustment of the whole reach.*

The local block reconstructed in SCAN B provides a basis to add information into the bundle adjustment of the entire reach attempted in SCAN A. This was achieved by pre identifying the low quality ground control points within the imagery to enable the identification of suitable image pairs within APS. This new alignment, entitled SCAN C1, solved the structure of the whole scene, matching 3871 cameras with 2,426,324 tiepoints providing continuous coverage from the Delta to the Fan reach. The resulting point cloud was then registered to a distributed network of 81 placed GCPs surveyed using a combination of RTK and PPK positioning based on a network of GNSS reference stations as described in Chapter 3 (section 3.3.1). This process of geo-registration in APS is undertaken as a rigid-body affine transformation (shift, rotate and scale), which results in the first model of the entire reach in an object space coordinate system.

While broadly continuous, this new scan incorporated a major discontinuity in the Gorge reach, associated with a step-like vertical offset of initially 30 m (see Figure 4.4) latterly reduced to 5 m (Figure 4.5).

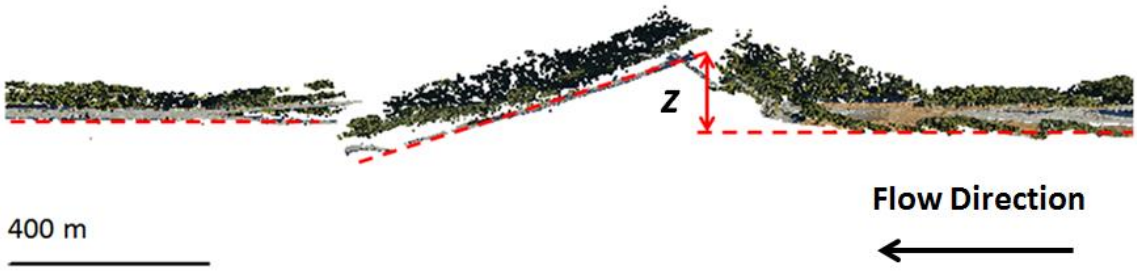

Figure 4.4 Offset within the sparse point cloud within the gorge of ~30 m, resulting in chunk separation and an incorrect geometry.

To reduce this offset a reconstruction was performed with additional exterior constraints, once more identifying known ground control (except this time across the entire catchment) within the imagery before the bundle adjustment.

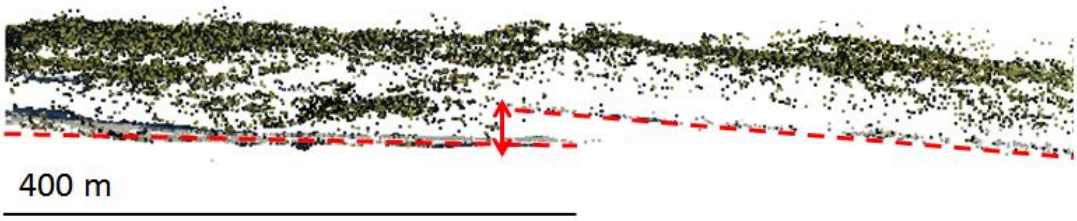

Figure 4.5 Improved geometry of reconstruction, however a vertical discontinuity of ~ 5 m can be seen to persist.

The degree of error was greatly reduced, however, as can be seen in Figure 4.5 a critical surface inconsistency remained. To address this a new bundle adjustment was solved, using additional exterior constraints in the form of taking the estimated 3D positions for the 3871 cameras solved previously and introducing these as constraints into a reconstruction.

With these additional constraints fixed, the reduced number of degrees of freedom lowered the overall computational overhead, enabling a solution to the scene structure to be found using the imagery at their native resolution in one single photogrammetric block. The resulting model, SCAN C2, comprised 3,747 matched cameras and 9.2 million tiepoints. This enhanced model had an average RPE for all tiepoints of 0.569 pixels in comparison to the 1.93 pixels for the smaller block, SCAN B.

4.2.2.4. *SCAN D: A Contiguous Sparse Model*

SCAN C2 incorporates considerable external constraints, some of which have been measured directly (e.g. the placed GCPs) others of which, are estimated from prior reconstructions (e.g. the camera positions). This new point cloud was refined iteratively to generate a final model. First, large outliers within the sparse cloud and their respective matches were visually inspected to determine if they were associated with a spurious feature match, or whether the anomalous position reflected an incorrectly

estimated camera position or pose. Any misaligned cameras were then reset and solved independently to predetermine their position while the rest of the scene structure was fixed. Any erroneous matches were also deleted.

The point cloud of tiepoints was further cleaned using a selection tool within APS, the gradual selection tool. This enables users to interactively select tiepoints that lie above or below a threshold value of RPE and/or reconstruction uncertainty. Focusing on reprojection error, which is largely related to incorrect feature matches (see discussion in section 3.4.4), the scene was successively realigned (BBA resolved) selecting only matches that were below a given threshold (starting with RPE = 1, then reducing the threshold to 0.9 and then 0.8 pixels). This iterative process reduces the density of tiepoint cloud, by incorporating only increasingly reliable matches and the associated exterior and interior camera geometry. Finally, the network of 81 GCPs is used to register the point cloud.

The final step in the workflow to generate a sparse reconstruction involves optimization of the bundle adjustment based on the GCP network. This step, described in section 3.4.4, involves local re-estimation of the exterior and interior camera orientation for cameras that incorporate matches tied to the GCPs. A series of local solutions to the camera geometry are then determined to estimate the optimal interior orientation (based on Brown's model of radial distortion, see Fryer, 1992), and then update the surrounding exterior orientation through the connected matching points. This process ties the final cloud of tiepoints tightly to the network of GCPs, resulting in low residual errors. However, it has the potential to create spurious compensation effects between the interior and exterior orientation that can give rise to significant systematic errors (see, James and Robson, 2014).

The final reconstruction comprised >15 M tiepoints at a mean density of 0.5 ppm^{-2}, a dense cloud of 340 million at a mean density of 10 $ppm^{-2}$and an average RPE of 0.6 pixels. Whilst there is full connectivity within the catchment topography, the Gorge Reach has significant local noise, indicative of poor reconstruction quality in this challenging section of the photogrammetric network.

4.3. Evaluating the 2014 model

4.3.1. GCP and ICP Residuals

The reconstructions of the 2014 survey were geo-referenced using a network of manually placed GCPs, comprising 1 m^2 heavy-duty, waterproof squares of red tarpaulin, pegged to the ground using 20 cm pegs, see section 3.3.2. The GCPs were deployed prior to the aerial survey and 3D coordinates obtained in NZTM using a combination of RTK and PPK GNSS observations. These observations were tied to a local network of 1st order reference stations which in turn were localized from the Land Information New Zealand (LINZ) Position NZ network of GNSS Continuously Operating Recording Stations (CORS); see Chapter 3 for further details.

The GCPs were distributed throughout the study reach in an attempt to provide quasi-uniform density and pattern of coverage. A spatially optimal deployment was, however, impossible due to: a) a requirement to site targets on stable areas of the floodplain that would not be disturbed by rising stage between deployment, survey and collection; b) restricted access to areas of private land; c) clear line of sight through marginal forest cover, particularly in areas of the Chinaman's Bluff and Gorge reaches.

It is important to recognize that residual errors determined by comparison between the model and GCPs incorporate a bias due to their role within multiple stages of the SfM workflows, i.e., as exterior constraints in the bundle adjustment; targets used to define a posterior registration of the sparse cloud; or as targets to optimize the exterior and interior orientation. Furthermore, the GCPs themselves are not without uncertainty and inevitably incorporate errors associated with the raw GNSS observations that include: occupation time; number of satellites and geometry of the constellation; timing errors; baseline bias; and errors in the location of the 1st order reference stations.

To provide a first assessment of model quality a simple leave-p-out cross validation test was performed in which 20 of the 81 GCPs were selected to provide independent check points (ICPs) while the remaining 61 targets were used in the registration and optimization analyses. The 20 check points were distributed across the study area at an approximately even density separated by a linear distance of typically 2-3 km (refer to Figure 4.6). A more formal approach to model testing using a Leave-One-Out-Cross-

Validation (LOOCV) would be preferable (essentially cycling the independent check data throughout space) however, the model reconstruction time (22 hours on the available computer hardware) provided a practical limit on the scope of sampling design at this scale.

Residual statistics for horizontal, vertical and total positional errors were calculated for reconstructions based on the network of 61 GCPs and 20 ICPs and then compared to models that incorporate all 81 GCPs in the solution. The results for both SCANs C2 and D discussed above are shown in Table 4.2.

Table 4.2 GCP and ICP residuals from both the initial (Scan C) and repeat (Scan D) complete reconstruction of the 2014 Dart; where ME is the arithmetic Mean Error; RMSE is the Root Mean Squared Error; SDE is the Standard Deviation of Errors; and MaxE is the Maximum Error is the maximum observed error. See Chapter 3 for a full derivation of all metrics.

GCP Configuration	Statistic	Total (m)	X (m)	Y (m)	Z (m)
	ME	1.161	0.106	-0.010	-0.037
SCAN C2	**MAE**	**1.161**	**0.476**	**0.554**	**0.691**
(20 ICPs)	SDE	0.613	0.633	0.643	0.947
	MaxE	3.011	1.604	1.439	2.525
	ME	1.002	0.036	0.017	0.066
SCAN C2	**MAE**	**1.002**	**0.371**	**0.390**	**0.694**
(81 GCPs)	SDE	0.710	0.506	0.484	1.006
	MaxE	5.292	1.520	1.702	5.173
	ME	0.441	0.082	-0.121	-0.105
SCAN D	**MAE**	**0.441**	**0.150**	**0.182**	**0.331**
(20 ICPs)	SDE	0.620	0.226	0.400	0.583
	MaxE	2.668	0.721	1.682	1.996
	ME	0.061	0.005	0.001	0.003
	MAE	**0.061**	**0.030**	**0.025**	**0.036**
SCAN D	SDE	0.098	0.056	0.039	0.092
(81 GCPs)	MaxE	0.622	0.385	0.171	0.583

A significant improvement in the GCP and validation statistics is clearly evident between SCANS C2 and D, measured in terms of both the ICP errors and also the full set of GCP errors. This reflects a progressive enhancement in the estimated camera geometry incorporated into each respective bundle adjustment. For example, SCAN C2 uses camera positions based on poorly constrained Google Earth GCPs derived in SCAN B and then refined in the first successful full block of the entire reach SCAN C1.

While these positions reflect a scene structure that has constructed a contiguous spatial model, the camera geometry is seeded from a low quality network of GCPs and these poor constraints result in the high observed residual errors. By contrast, SCAN D takes as its starting point camera reference positions from SCAN C2 and then refines these

further by recalculating the scene geometry and then optimizing this using the accurate GCP network. The result of this iterative process is to more than halve the independently estimated (ICP) vertical (0.691 to 0.331 m) and total 3D errors (1.161 to 0.441 m). The variance of the residuals (measured in terms of the SDE) is significantly reduced in the vertical (0.947 to 0.583 m) between these two scans.

Overall, the residual errors appear to indicate little structural bias at the scale of the entire model (i.e., a consistent horizontal or vertical offset). This is revealed by the low mean error observations (ME), for either of SCANS C2 or D and either model of constraints (61 or 81 GCPs) which in almost all cases are below 0.1 m in any direction. The most significant component of the 3D error is consistently the vertical dimension. For the model based on 61 GCPs and errors measured with respect the 20 ICPs, the vertical MAE varies between 0.691 m and 0.331 m for SCANS C2 and D, respectively, compared to horizontal errors of between 0.476 m (x) and 0.554 m (y) for SCAN C2 and 0.150 m (x) and 0.182 m (y) for SCAN D. The internal variance of the residuals, as revealed by the SDE, is between 30-50% higher in the vertical than horizontal axes.

The presence of significant outliers in the residuals is highlighted by the maximum error metric (MaxE). In the SCAN C2, based on the full 81 GCPs, the 3D maximum error is 5.267 m. This high residual is associated with a marker between the Hillocks and Paradise (Figure 4.6, shown below). For comparison, Figure 4.7 displays the corresponding residuals for SCAN D.

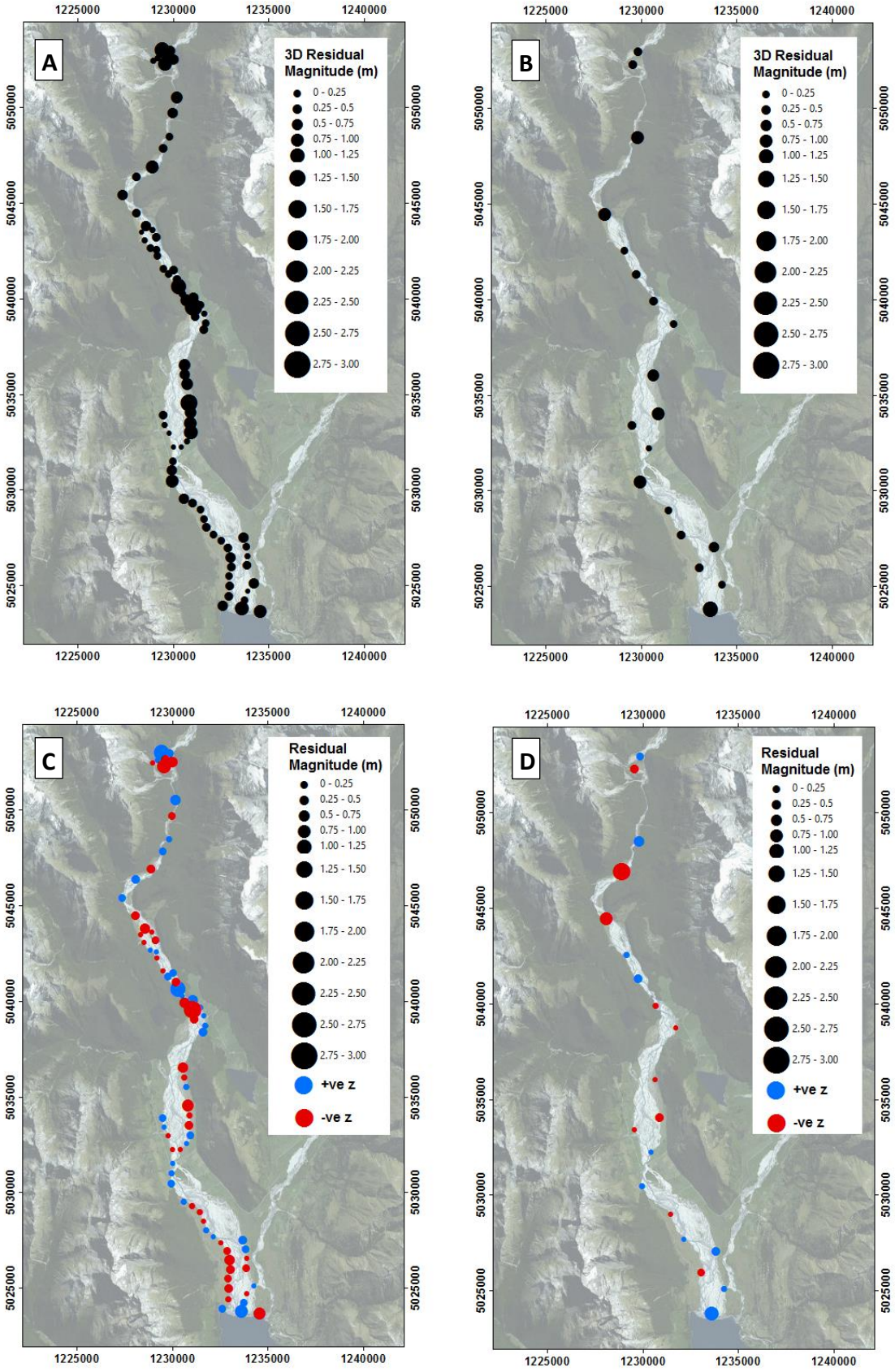

Figure 4.6 Spatial distribution of GCP and ICP errors within the study Area. A) 3D GCP error, B) 3D ICP error, C) Vertical GCP error, D) Vertical ICP error. *Projection: NZTM 2000*

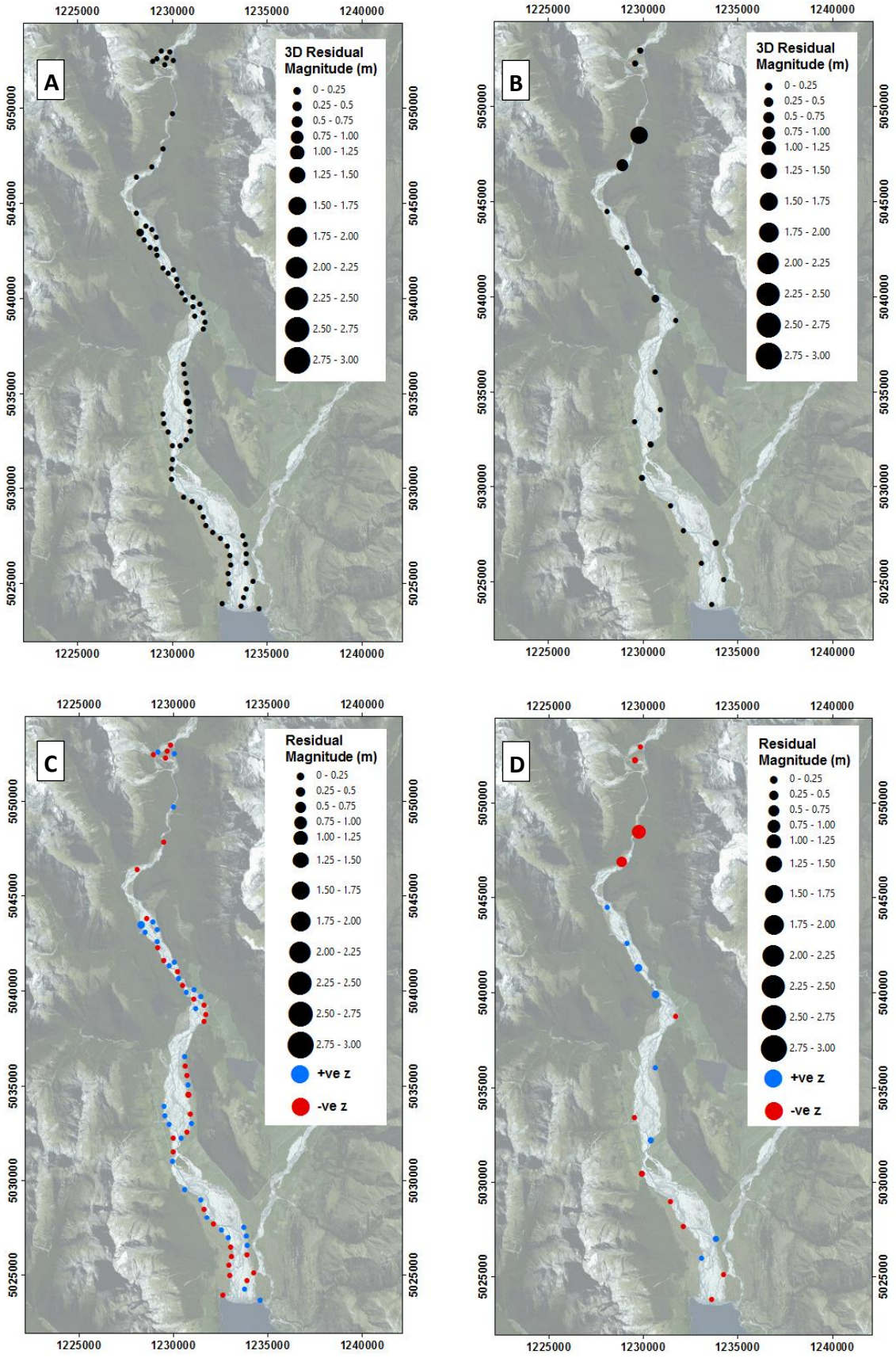

Figure 4.7 Spatial distribution of GCP and ICP errors within the study Area. A) 3D GCP error, B) 3D ICP error, C) Vertical GCP error, D) Vertical ICP error. *Projection: NZTM 2000*

The high magnitude residuals in SCAN C2 are mitigated effectively by the processing steps leading to SCAN D. In this case, the maximum residual error is reduced to 2.668 m for the independent test data and 0.622 m for the experiment in which all GCPs are used to register and optimize the model.

It is useful to contextualise the results in terms of typical SfM performance and the maximum theoretical level of accuracy that could potentially be achieved. Luhmann (2003) suggests that the maximum theoretical accuracy achievable photogrammetrically can be estimated as a function of the flying height, baseline between image pairs and the camera focal length:

$$S_z = \frac{h}{b} \times \frac{h}{c} S_{px} \qquad [4.1]$$

Where, Sz is the accuracy in the camera direction (m), b is the baseline distance (m) between image pairs, c is the focal length (m) and S_{px} is the image precision, taken as the dimensions of the CCD elements (m). While there is some inevitable variability in the baseline distance between image pairs and the flying height in the manually flown flight plan, taking typical values of 85 m for the baseline separation (2 s image frequency at 85 knots); 350 m for the above ground height; a fixed 0.028 m focal length; and image precision of 5.503 μm, suggests a theoretical accuracy of 0.27 m. This measure applies to reconstructions based on stereo pairs. In this application, the 'reference' setting used in APS derive matches from multiple pairwise observations that are then refined as a set. This makes direct application of Equation 4.1 somewhat complicated but it is broadly equivalent to the characteristic maximum accuracy of 1/1000 of the flying height described by James and Robson (2012). More recently ,Smith and Vericat (2015) summarise empirical quality estimates of SfM models derived from mainly from imagery derived from Unmanned Aerial Systems (UASs) flown at slower velocities and from lower flying heights. In a summary of over 50 case studies and using the RMSE as a measure of model quality, they find applications tend to achieve a somewhat lower level of accuracy, on average approximately 1/639 of the flying height. Taking these three estimates of potential model quality; 0.27 m (Luhmann, 2003); 0.35 m (James and Robson, 2014); 0.55 m (Smith and Vericat, 2015), it is clear that the MAE of the ICPs in

SCAN D of 0.441 lies somewhat above the maximum theoretical accuracy but favourable with comparable studies.

These statistical measures of model quality, while useful to provide a broad indication of model fit, do not provide a detailed insight into any spatial systematic errors that might be present in the 3D sparse point cloud. As discussed in Chapter 2, SfM reconstructions have been found previously to be sensitive to the incorporation of systematic errors that emerge from the simultaneous solution of the exterior and interior orientation parameters in the self-calibrating bundle adjustment (Javernick et al., 2014; James and Robson, 2014). These effects are evidenced by the presence of dome or bowl distortions that span the photogrammetric block and reflect errors in the estimated radial distortion parameters of the Brown-Conrady model used to represent the interior camera orientation.

In order to explore the presence of such structural effects, the spatial distribution of model residuals is plotted previously in Figures 4.6 and 4.7 for SCAN C2 and SCAN D respectively. Residuals are shown here as proportional symbols that are scaled relative to the magnitude of the total 3D and vertical errors respectively. Results are shown for both residuals associated with a model incorporating 81 GCPs and then one based on 61 GCPs and a set of 20 ICPs. These figures reveal significant spatial autocorrelation of residuals, most evident in the plot of GCP residuals due to the higher spatial density. For example, in both SCAN C2 and D there are clusters of high 3D error magnitudes in the Fan reach and then downstream in the lower end of the Chinaman's Bluff reach and also along the true right of the Paradise reach. By contrast, low GCP residuals are evident in both scans throughout the delta Reach.

Decomposing the residuals into their horizontal and vertical components provides an insight into the direction as well as the magnitude of errors. For example, the vertical error calculated for the independent check points (ICP) in SCAN D shows an oscillating change in the direction of the residuals. In these plots (Figures 4.6c/d and 4.7c/d) negative residuals (model below GCP) are plotted in red and positive residuals plotted in blue. In Figure 4.7b and 4.7d the ICP errors exhibit a characteristic switching of direction, negative in the Fan and Gorge reach, then positive through the Chinaman's Bluff reach and so on. Despite this apparent spatial structure, there is little convincing

evidence of the bowl or dome type errors that arise from incorrect calibration of the radial distortion and focal length parameters of Brown's model of lens distortion.

To test for the presence of dome/bowl errors directly, Figure 4.8 shows the pattern of residuals with distance from the centre of the GCP network. Doming errors, if present, would be expected to reveal a systematic increase away from this centre point, something that is not evident for either SCAN C2 or D and neither for the 3D or vertical errors. The absence of such a systematic correlation highlights the key role that the final process of optimization plays in mitigating these systematic errors. This step uses ground control to resolve the bundle adjustment at each GCP and then propagates the revised orientation parameters throughout the model. It is noteworthy, however, that the most significant ICP errors occur within the Gorge Reach and at the extremes of the model domain, the Fan and the Delta Reaches.

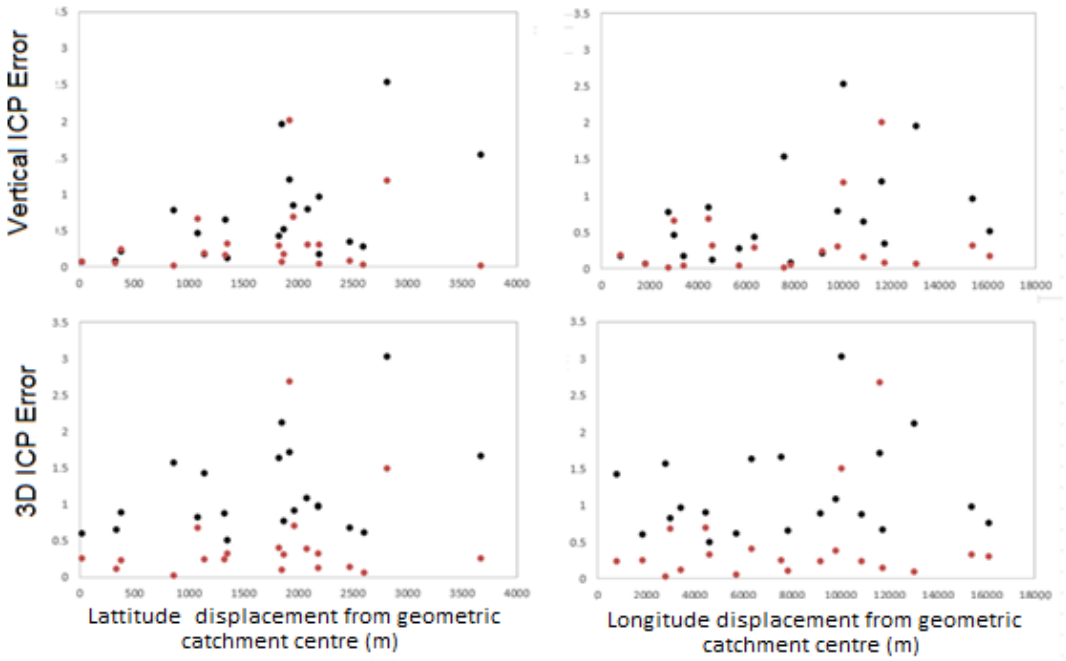

Figure 4.8 ICP values for Z and 3D displacement of markers plotted against the distance from the GCP centroid location for Scan C (Black) and Scan D (Red).

4.3.2. The Effect of Optimization of the Model Structure and Systematic Errors

The use of the GCPs as exterior constraints to optimize the model locally would be expected to dramatically improve the model fit at these control points. It is less clear, however, whether this will result in a general improvement in the model, or whether this could in turn add further systematic bias, particularly if there are errors associated with the GCP positional data (i.e., local blunders) or existing systematic errors in the prior exterior camera geometry.

A useful insight into the effect of optimization is to compare the geometry of the model before and after this key step in the processing chain. To facilitate this analysis, the sparse clouds of tiepoints for both SCANS C2 and D were extracted both before and after the final optimization step. Each point cloud was then decimated using the Topographic Point Cloud Analysis Toolkit or ToPCAT (Brasington et al., 2012) to facilitate the comparison of the models using raster differencing methods. ToPCAT was used to sample the point cloud and return the mean elevation in 5 x 5 m windows, along with a statistical summary of this local cloud. The reduced resolution dataset was then interpolated using Delaunay Triangulation to derive a Triangular Irregular Network (TIN) that was then linearly resampled to a raster DEM at 5 m resolution. The before and after optimization DEMs were then compared by subtracting the posterior model from the prior to create a DEM of Differences or DoD. The results for the two scans are shown below in Figure 4.9; note that strongly contrasting colour-ramps are used to scale the resulting maps of difference.

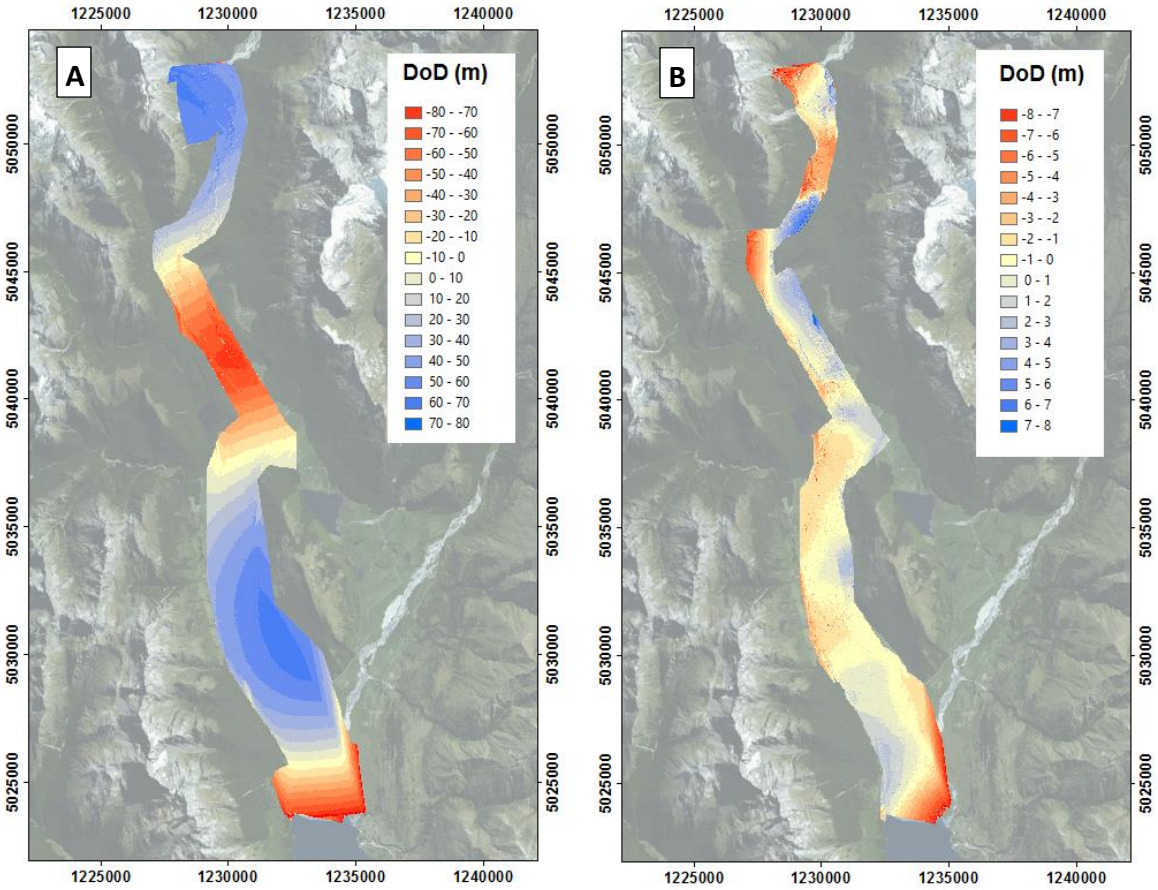

Figure 4.9 Optimization DoD, displaying the surface impacts upon reconstruction geometry during the camera calibration refinement stage. A) Scan C, B) Scan D. *Projection: NZTM 2000*

The resulting pattern shows the spatial correction calculated and applied during the optimization process. The difference in the magnitude of adjustment between SCAN C2 and D is striking, varying by an order of magnitude with typical adjustments of tens of metres (and up to ±80 m) occurring in SCAN C2. Almost as striking is the clear systematic, spatially correlated nature of the correction applied, which is also most marked in SCAN C2. In this scan, optimization leads to a wave-like pattern of correction, with positive adjustment in the Fan and Gorge reaches alternating into negative adjustments through the Chinaman's Bluff reach, before finally switching back to positive adjustment and then negative once more through Paradise-Hillocks and the lower Delta Reach.

The optimization of SCAN D is much less pronounced, but nonetheless still exhibits strong systematic variability and oscillating patterns of positive and negative vertical

adjustment. An indication of this systematic pattern of correction is also evident in the distribution of vertical ICP residuals shown in Figures 4.6 and 4.7 above. This suggests that even within areas of the GCP network, the optimization process is correcting a regional systematic pattern that is magnified significantly beyond the spatial perimeter of ground control. While there is no reference data to corroborate this, it seems likely that the pattern of distortion will quickly become aberrant and render these areas unfit for further analysis.

Intriguingly, while previous studies have also shown systematic distortion for SfM point clouds (e.g., Javernick et al., 2014; James and Robson, 2014) this has typically taken the form of a single dome or bowl distortion that extends across the entire photogrammetric block. By contrast, even in the better constrained SCAN D, the systematic adjustment applied during optimization is acting to correct a much more complicated spatial pattern that appears to reflect multiple regions of contrasting adjustment. This pattern may partly be explained by the effects of parameter compensation that occur during the bundle adjustment and later optimization. The optimization process solves the optimal interior geometry for the scene which is applied across all cameras. Under conditions where prior errors in exterior camera geometry and distribution of ground control are spatially uniform, inaccuracies in the interior camera model dominate the resulting systematic spatial structure errors in the 3D point cloud. The classic dome or bowl shaped error then reflects the principal sensitivity of the K1 radial distortion parameter and its compensation with the estimated focal length.

By contrast in this study, there are strong variations in: a) the density of the camera network; and b) the distribution of ground control. Perhaps more important still, however, is that in order to generate a solution for the entire reach as a single block, it was necessary to seed the bundle adjustment with localized and low quality exterior constraints, specifically, the Google Earth derived GCPs. These were then used to provide first order estimates of the exterior camera positions for SCAN B, which were themselves used to provide exterior constraints for SCAN C1 and, in turn, these were used as constraints in SCAN C2 and so on. Iterating the bundle adjustment in this manner enabled a solution to the structure of the whole reach, but in so doing potentially incorporated systematic bias at the start of this progressive process. These

incorporated errors can then be compensated only partially in the final optimization process where the network of GCPs is used as a set of exterior constraints. An important consequence of this is that the optimization process is forced to generate strong, spatially distributed adjustments in the exterior geometry of the model which are evident Figure 4.10. This shows the change between the pre- and post-optimization exterior camera positions and clearly demonstrates the strong spatially auto correlated adjustments that match the difference models shown above in Figure 4.9.

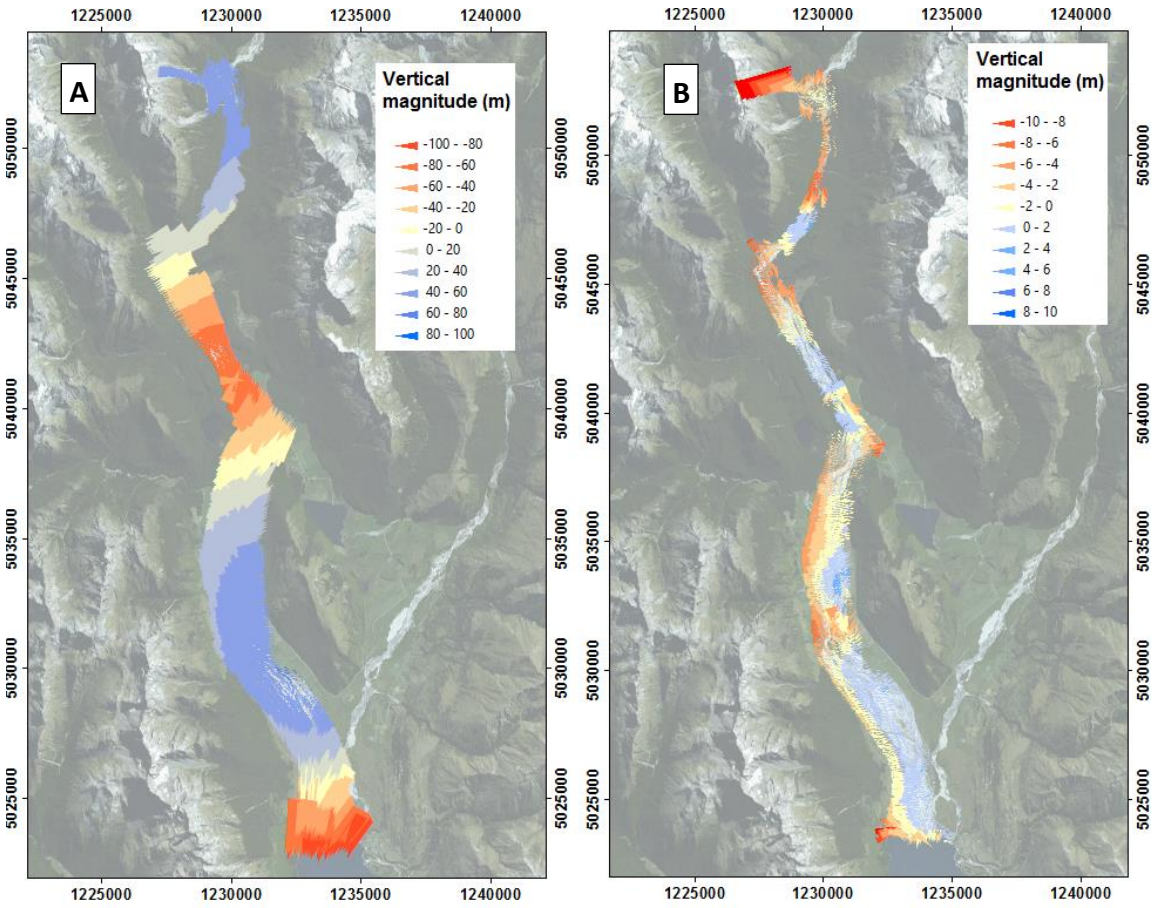

Figure 4.10 Spatial distribution of changes in camera positions (3D and vertical) pre and post-optimization for Scan C2 and Scan D. *Projection: NZTM 2000*

The spatial distribution and magnitude of camera change mimic the surface changes seen during the optimization phase. The average camera positional change during optimization of Scan C2 is 61.647 m, whereby this value is reduced to just 3.130 m for Scan D. The impact of such large extrinsic variation may be seen through tracking the polynomial intrinsic parameters responsible for estimating radial distortion of the lens, K1, K2 and K3. Comparing the percentage change in each parameter between C2 and D

during the optimization phase, K1 (0.227% vs. 0.0155%), K2 (-0.521% vs. 0.268%) and K3 (-8.606% vs. -1.458%), shows a change in the primary radial distortion parameter of an order of 14 times greater. Despite these significant calibration adjustments, the final average camera positional change between the two scans is just 2.137 m. This reinforces the notion that errors in the exterior pose accumulate within the bundle adjustment via compensation in the intrinsic calibration process.

4.3.3. Benchmarking the Reach Scale Model on Existing Data

The analysis of GCP and ICP errors provides a very partial sampling of the resulting 3D products from the SfM reconstruction. In an ideal scenario, the resulting point cloud and associated terrain products would be benchmarked against a higher quality reference dataset in the form of continuous DEM or point cloud. Unfortunately, there are no available existing high resolution terrain data products for the Dart catchment and the highest resolution available DEM is an 8 m resolution model derived from 1:50,000 mapping with 20 m contour intervals. Moreover, the area of interest is dominated by highly labile fluvial topography and significant landsliding and debris flow activity, so that any historical data are likely to reflect significant differences at the scale of the topographic features of interest in this thesis.

A useful insight into the detailed structure of model quality can however, be provided by comparing the models generated to an existing survey of the lower 5 km of the Delta Reach through an airborne LiDAR survey commissioned by the Queenstown Lake District Council in 2011 (see section 3.5.1). For this survey, a 1 m DEM based on last pulse data was made available (Figure 4.11). This survey extends over the active valley floor and is dominated by dynamic fluvial topography and significant areas of wetted channels and has not been corrected to remove wetted channels or process them to incorporate bathymetry.

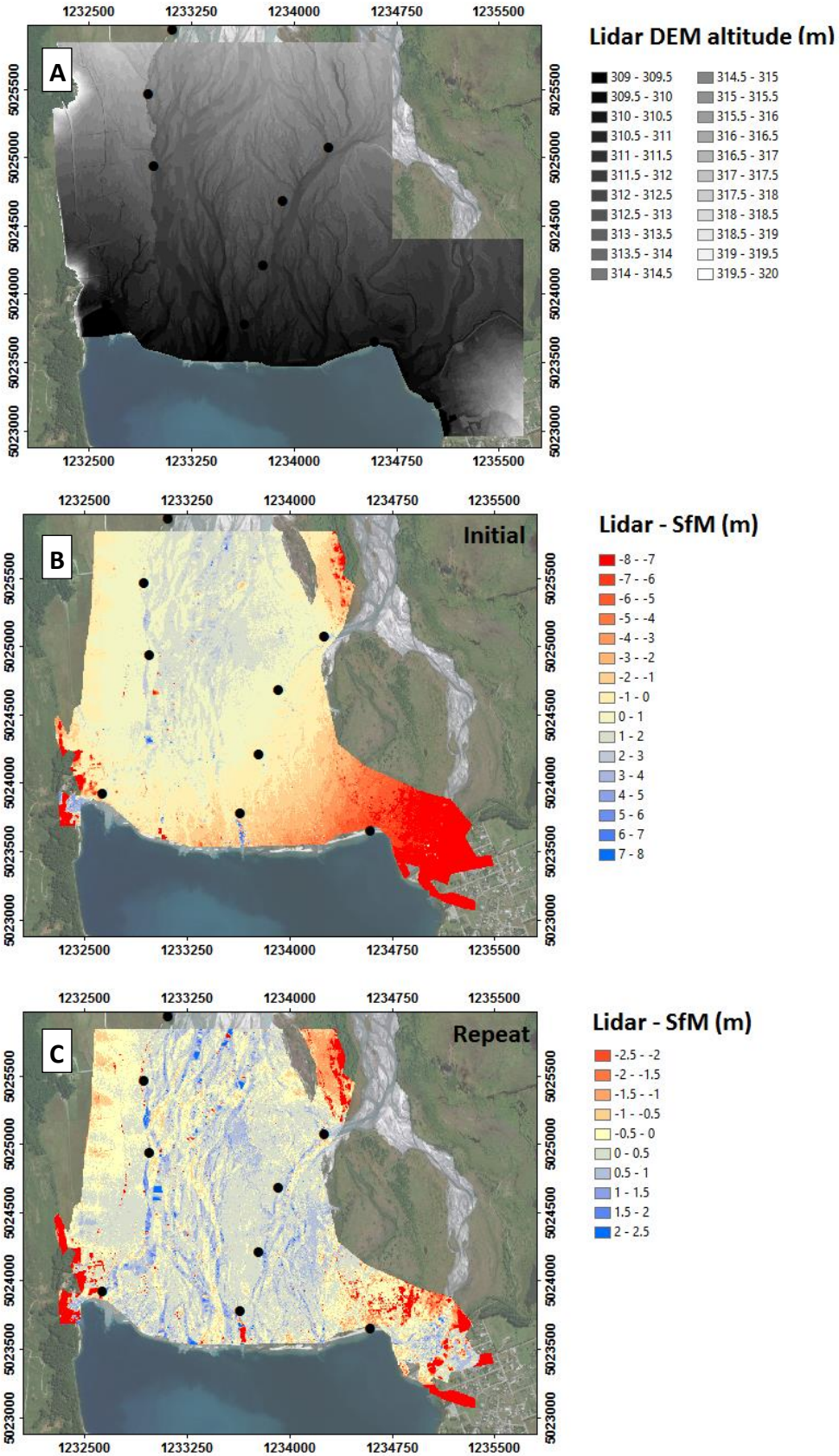

Figure 4.11 A) Otago Regional Council LiDAR DEM, B) LiDAR-SfM SCANC 2 and C) LiDAR-SfM SCAN D DoD. As seen previously the scales between the two scans are different. *Projection: NZTM 2000*

At this stage of processing the 2014 point clouds have also not been corrected to either mask or remove wetted areas, so any comparison between these two datasets will inevitably incorporate significant areas that have either changed topographically and/or are wetted. Nonetheless, DEM differencing between Scan C2 and Scan D and the LiDAR model reveals some clear differences between the two stages of the SfM processing. To facilitate comparison of these datasets, the sparse clouds for the two scans were processed using ToPCAT as described above, to generate 2 m resolution DEMs based on the arithmetic average elevation.

The DoD for SCAN C2 shows that within the central area of the model covering the braidplain, differences between the two models are typically within $\pm$1 m vertically. However, as predicted, once outside of the control network, systematic errors result in significant distortion of the SCAN C2, with increasingly negative deviations between the LiDAR and SfM model (i.e., LiDAR is lower) evident in the lower true left of the Delta and over the township of Glenorchy. Here, the pattern of distortion appears to reflect the classic radial distortion bowl shape, with similar deviations occurring at the upstream of the comparison near the confluence with the Rees River on the true left.

By contrast, the DoD based on SCAN D exhibits a much tighter fit to the LiDAR survey, 71% of surface error within $\pm$0.5 m. The systematic errors over Glenorchy are less evident and deviations in this area are of the order of 1 m and up to 2 m near the upstream Rees-Dart confluence. In this DoD, the dominant spatial structure relates to the observed pattern of channel change between 2011 and 2014, and reveals extensive reworking of the channel bed, particularly on the true right of the delta.

4.3.4. Summary of model performance

Results from the catchment wide survey highlight two key findings: (1) GCP residuals are improved when minimal change is required during the optimization phase, which adjusts parameters (interior and exterior) based upon the GCPs. A perfect reconstruction can be assumed to, therefore, exhibit no difference pre- and post-optimization. This may be achieved through accurate camera location estimation (exterior orientation) and through improved radial distortion estimation (interior orientation); (2) whilst camera reference data can be used to relax the computational

overhead of the alignment process, there is a limit to the effective tolerance of the positional errors in the input data. In the case of the initial reconstruction generated from a combination of informal GCPs within the gorge and a few additional at the periphery, large errors in the exterior constraints were propagated into the bundle adjustment. These errors can be diminished by using updated camera positions, back calculated from a GCP informed scan, however, the magnitude of camera errors is still of the order of mean = 3.130 m. By comparison, typical GPS systems operate with an error of +/- 10 m and it is important to assess if this value is sufficient to minimize the propagation of further errors.

Finally, while LiDAR data were available to provide a detailed spatial analysis of the resulting model, the time elapsed between the surveys and the partial coverage of this model complicate a simple comparison. To address this final issue, a TLS survey of the debris fan was undertaken to provide an up-to-date reference data set for comparison.

4.4. High Quality Control Site

4.4.1. Rationale

As seen previously, there is a requirement to provide a more up-to-date comparison dataset to provide a spatially continuous representation of error within the 2014 SfM survey. A TLS generated point cloud of the Slipstream debris fan provides a means to assess how imagery captured under catchment wide SfM survey conditions performs at a local scale and further test strategies to improve the workflow during processing.

The presence of a control dataset provides a basis from which to assess the impact of survey refinements upon SfM reconstruction quality. The literature has indicated the potential to improve results through optimizing image selection (Furukawa and Ponce, 2010; Niethammer et al., 2012; Westoby et al., 2012), pre-calibration strategies (Bertin et al., 2015; Gneeniss et al., 2015) and introducing additional exterior constraints, both pose estimation and ground control (Smith et al., 2016) and ground control (James and Robson, 2012; Turner et al., 2012). The following experiments aim to assess the potential for such improvements and how they may be relevant to a landscape scale SfM survey, specifically within APS.

4.4.2. Slipstream Debris Fan

The debris fan at Slipstream, located in the most upstream study reach, was chosen as a control site to provide 3D spatial analysis of SfM reconstruction. The debris fan surface was presumed to be constant across the period of data capture and provided a predominantly dry surface of sufficiently variable topography with relatively dense ground control (Figure 4.12). A benchmark survey was conducted over a two day period from a Reigl long range TLS scanner (of range 1400-1000 m) located at 7 discrete locations across the surface of the fan to provide complete coverage of the area (see section 3.5). Registration of the TLS data was achieved with 21 GNSS positioned targets, giving a mean absolute error of -0.103 m.

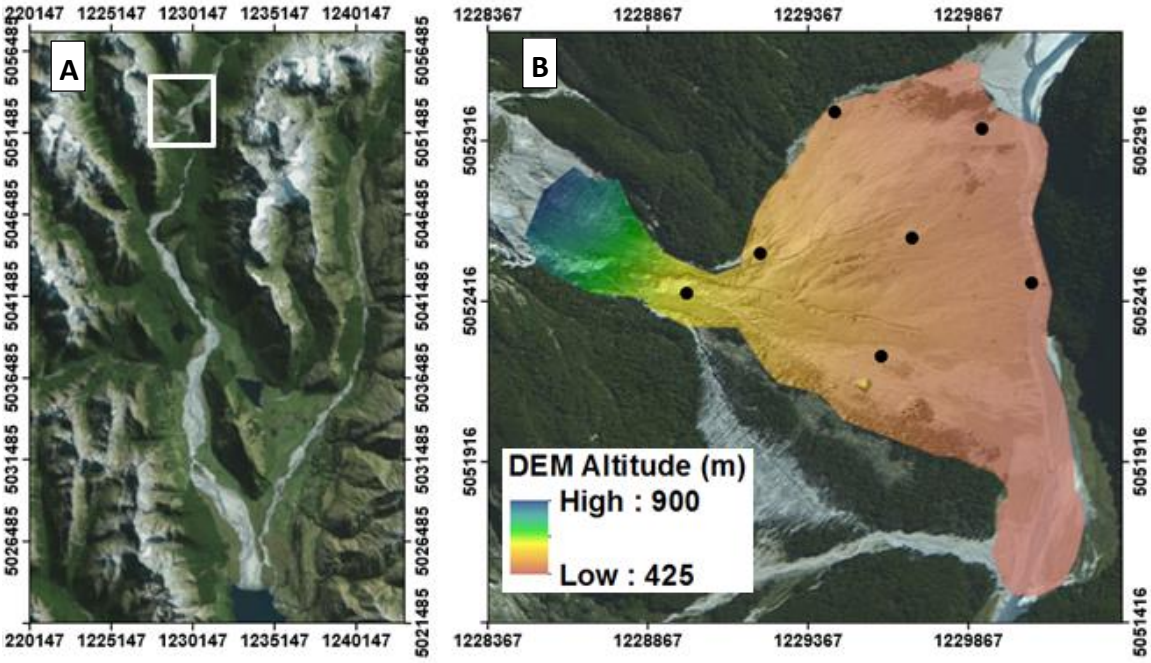

Figure 4.12 (A) Relative location of the Slipstream Debris fan within the Dart catchment and (B) aerial imagery with altitude projection. *Projection: NZTM 2000*

For comparison, a total of 469 images were selected for the construction of a new photogrammetric block and referenced via 7 GCP targets distributed across the fan surface. The following experiments are presented in part to further de-mystify the SfM pipeline within APS and to derive information that may inform subsequent survey campaigns. A number of inferences have been made when relating the processing steps

undertaken in APS (see section 3.4) to the generalised SfM pipeline presented in Chapter 2.

4.4.3. Experimental Design

In an attempt to interrogate the SfM pipeline within APS, four experiments were conducted, predominantly aimed at addressing the impact of tiepoint quality, interior and exterior constraints. This was achieved through addressing the impact of:

4.4.3.1. ***Image Quality:*** The image quality within SfM relates to both the sensor quality and scene subject, with improvements thought to eliminate spurious matches that may influence the bundle adjustment resulting in deterioration in model quality (Furukawa and Ponce, 2010). In line with standard SfM procedures, problem images are removed through visual inspection, excluding images with >90% water or vegetation, which are known to provide mismatches (James and Robson, 2012; Niethammer et al., 2012; Westoby et al., 2012; Fonstad et al., 2013). A further step was taken to remove images identified as misaligning across repeat identical reconstructions by a magnitude greater than the mean camera error plus two standard deviations. This is simply achieved by exporting and cross referencing predicted camera positions from three identically run reconstructions and identifying inconsistent alignments. The remaining 350 images are deemed to consist of sufficient detail to reduce keypoint mismatching.

4.4.3.2. ***Pre-Calibration (Cal0; Cal3):*** Introducing internal constraints is thought to minimize the identification of an incorrect local minima during the bundle adjustment (Gneeniss et al., 2015). Geo-referenced and optimized scans generate a camera calibration file as part of the model construction within APS. These calibration parameters can be exported and used as the initial boundary conditions for a new SfM reconstruction. This process may be applied iteratively until results deteriorate. The pre-calibration reconstructions were undertaken with the use of the reduced imagery set defined by the image quality experiments (4.4.3.1). This is a

result of serially misaligning reconstructions and model splicing occurring when iterative calibration was attempted with the original imagery set.

4.4.3.3. ***Inferred Camera Pose:*** *A priori* camera positional data have the potential to greatly reduce computational time within the SfM pipeline as matches are sought only from spatially coincident images as opposed to the entire image set. Section 4.2 identified that catchment scale reconstructions require the use of *a priori* camera positional data to successfully build reconstructions. It is important, however, to assess how this compares to the 'Disabled' pairwise-selection setting available in APS, in which no constraints are introduced. Smith et al. (2016) highlight the potential for high quality pose data to minimize requirements for ground control, however, the quality of pose estimation from standard GPS units potentially introduces additional error into the reconstruction. In the absence of GPS data for 2014, existing scans were used to export camera location information to act as an equivalent, to test the accuracy and time taken to generate reconstructions under the Reference settings. Exported co-ordinates are significantly more accurate than standard camera GPS are able to achieve, necessitating a series of reconstructions to test the robustness to error through randomly distributed x, y, z error of up to 10 m. The introduction of exterior constraints may be assessed alongside the use of pre-calibrated cameras (section 2.2.2) to determine if this: a) improves results further; and b) reduces computational time.

4.4.3.4. ***Additional Ground Control:*** James and Robson (2012) highlighted the importance of ground control extent, with significant deterioration in reconstruction quality beyond this. The lack of ground control in the southern toe of the debris fan is rectified through the extraction of additional control points from the TLS point cloud. A previously corrupted GCP (Checkpoint 2) and an additional, easily identified branch (Checkpoint 1) on the fan surface were chosen (these serve as ICPs for the prior experiments). The TLS and SfM point clouds were loaded within Cloud Compare and fitted to each other using an Iterative Closest Point

algorithm. The identifiable GCPs were then selected and the displacement between the SfM and TLS measured. These new GCPs are used to assess the impact of ground control extent and how this relates to the optimization process within APS.

4.4.4. Experiment Implementation

The results of each experiment may be seen through a DoD comparison with the TLS control data obtained for the Slipstream debris fan (Figure 4.13; 4.14; 4.15; 4.16; 4.17), while the relevant error metrics are presented in table 4.3. The Geomorphic Change Detection (GCD) ARC plugin (Wheaton, 2010) was used to perform change detection calculations and derive respective histograms that describe the distribution of differences.

To eliminate outliers associated with spurious tiepoints and holes in the datasets the reclassify tool was first used to remove all regions outside the range of +/- 5 m. Across the majority of experiment reconstructions, a significant block of error of >2 m is present within the GCP radius, proximal to the helicopter take-off and landing where image overlap is insufficient resulting in a hole in the mesh. While the hole itself is removed from the analysis, it is interesting to note the ring of increased error resulting from a reduction in image density at that location.

Table 4.3 Key stats for each scan tested. Mean error is a representation of inherent bias. The additional GCPs derived from the TLS were used as Checkpoints 1 and 2 and therefore fail to provide much information regarding the final Extra GCP reconstruction.

Scan	Sparse Points	LOOCV (m)	Mean Error (m)	% < 0.5 m	Standard Deviation (m)	Check point 1 error (m)	Check point 2 error (m)	Time (Minutes)
Standard	1,126,181	0.303	0.023	72.27	0.875	0.741	2.605	525
Image Quality	1,106,185	0.312	0.026	71.50	0.924	0.561	2.468	325
Pre-Calibration	1,109,353	0.315	0.104	74.62	0.802	0.329	2.013	315
Inferred Pose	1,128,412	0.303	0.055	72.80	0.855	0.507	2.420	35
Extra GCP	1,126,181	0.560	0.117	77.40	0.792	*0.064*	*0.031*	525

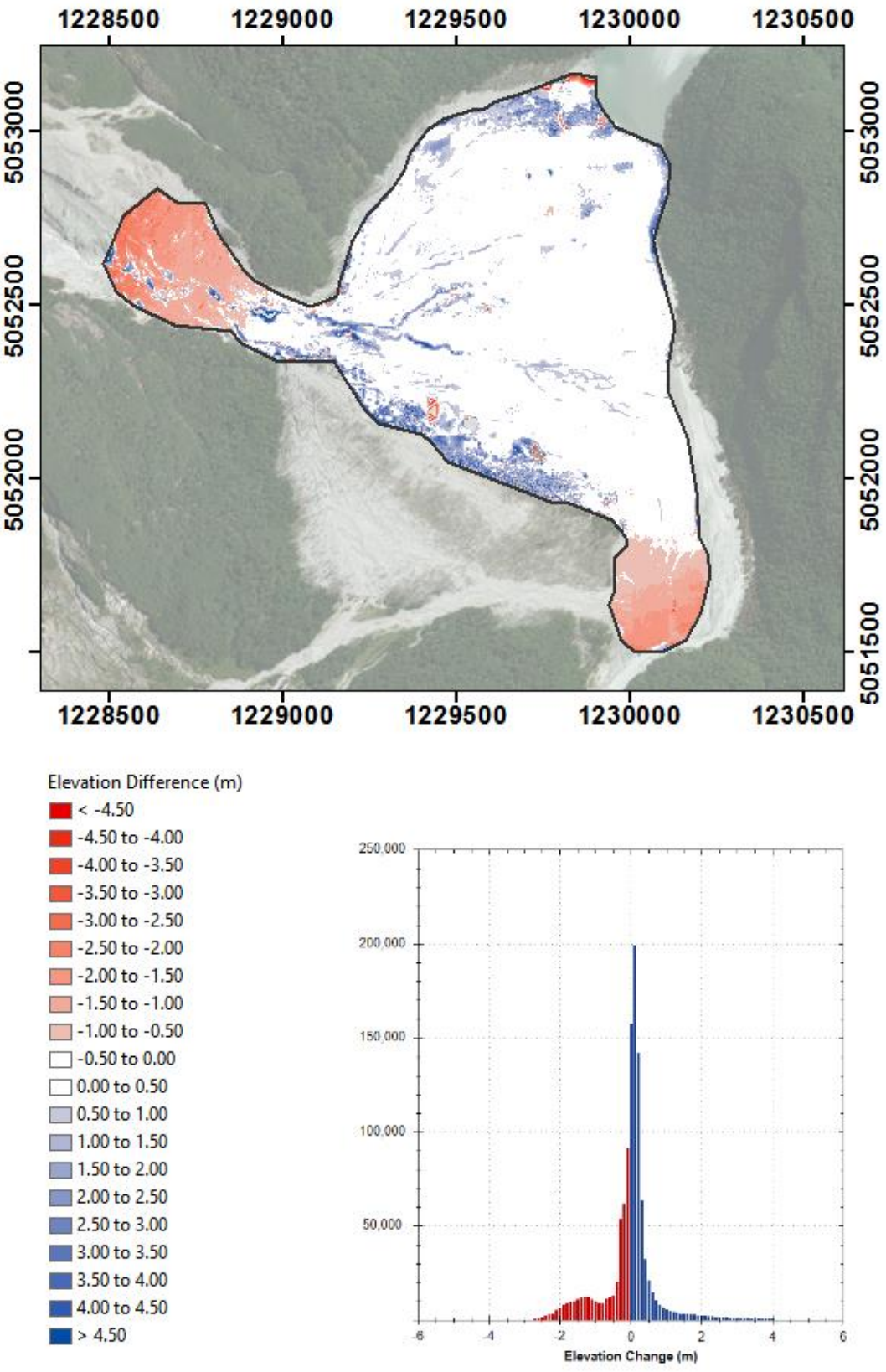

Figure 4.13 TLS-SfM DoD obtained under a standard SfM workflow within APS, see 4.4.3.1. Significant distortion is observed at the steep section of the fall and the Southern toe. *Projection: NZTM 2000*

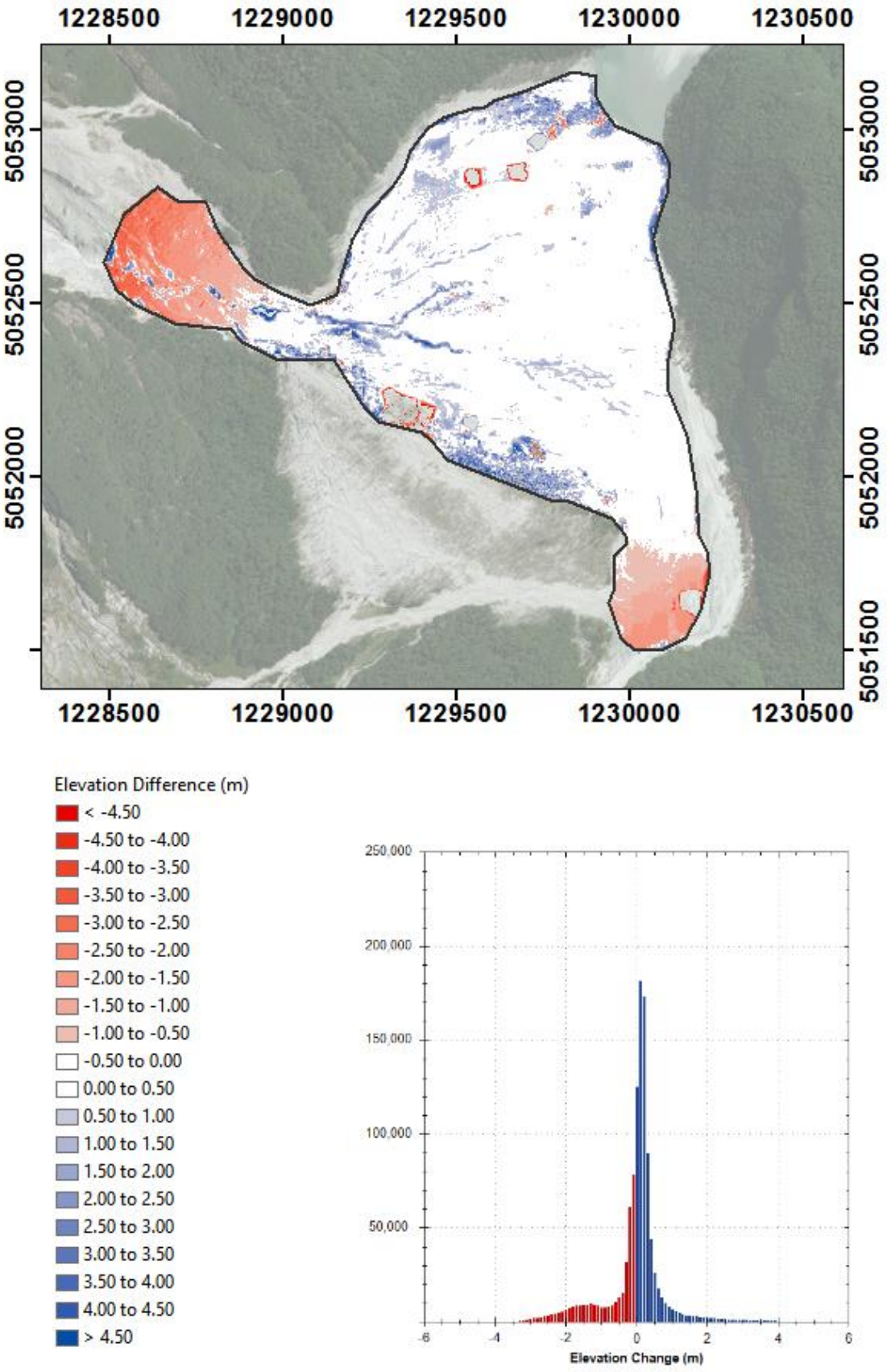

Figure 4.14 TLS-SfM DoD obtained under a standard SfM workflow with the removal of low quality imagery within APS, see 4.4.3.2. There is a slight reduction in distortion at the toe, however, a number of hole have appeared. *Projection: NZTM 2000*

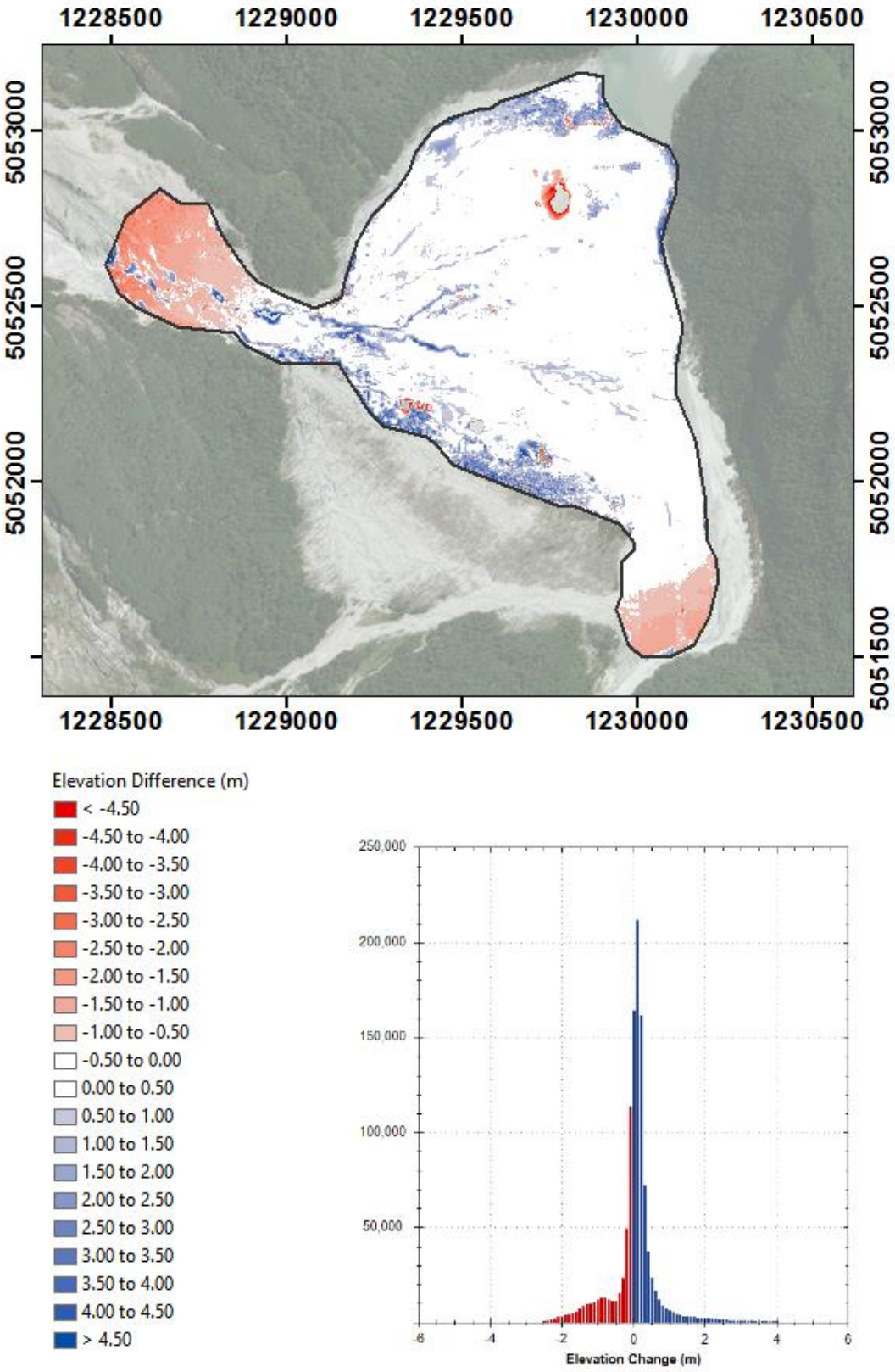

Figure 4.15 TLS-SfM DoD obtained under the optimal pre-calibrated setting within APS. Projection, see 4.4.3.3.Further reductions are observed at the toe as well as the steep section of the fall. *Projection: NZTM 2000*

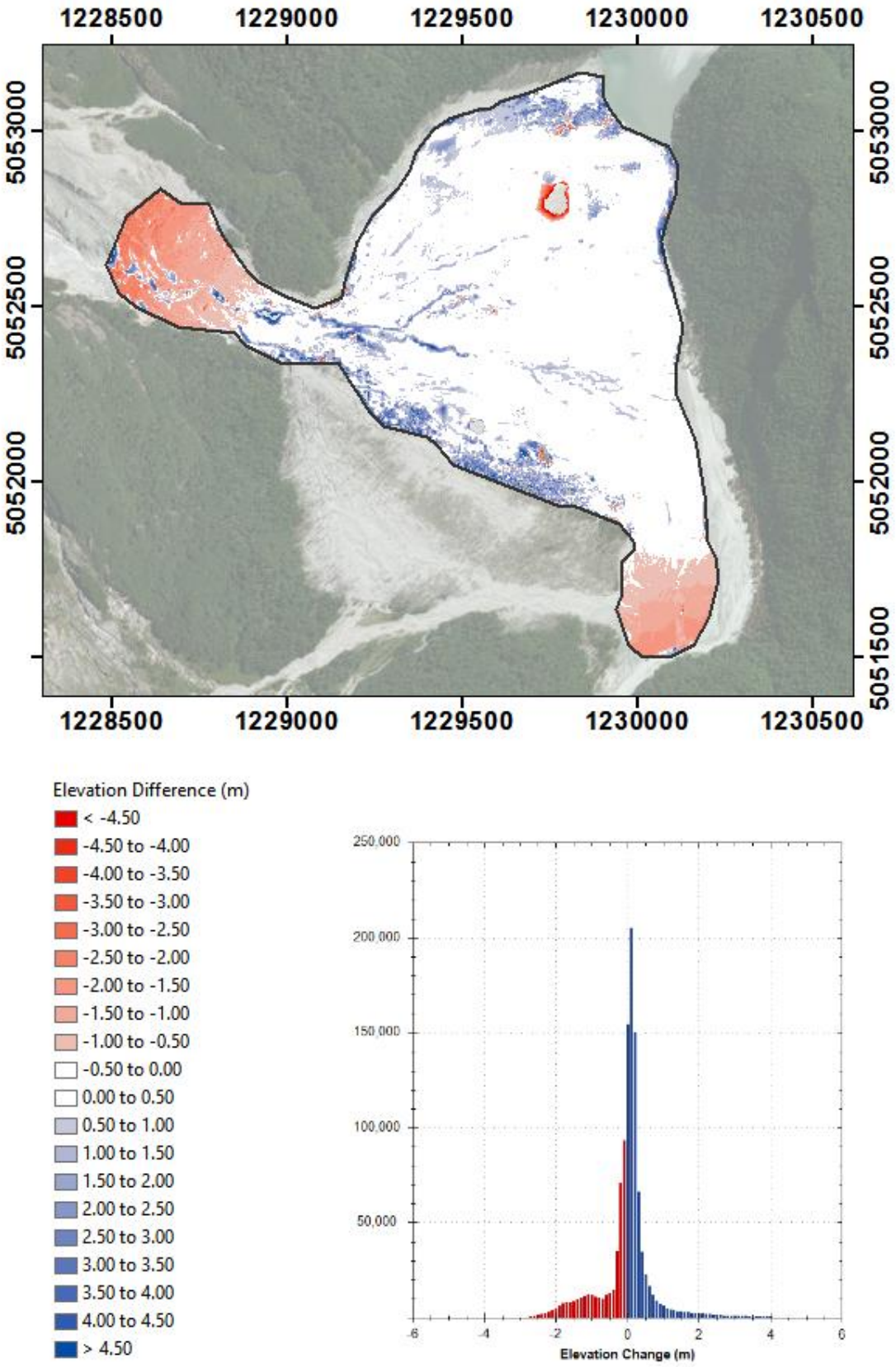

Figure 4.16 TLS-SfM DoD obtained under inferred camera positions within APS, see 4.4.3.4. Results show a marginal improvement of surface geometry, however, significant time gains are made during computation. *Projection: NZTM 2000*

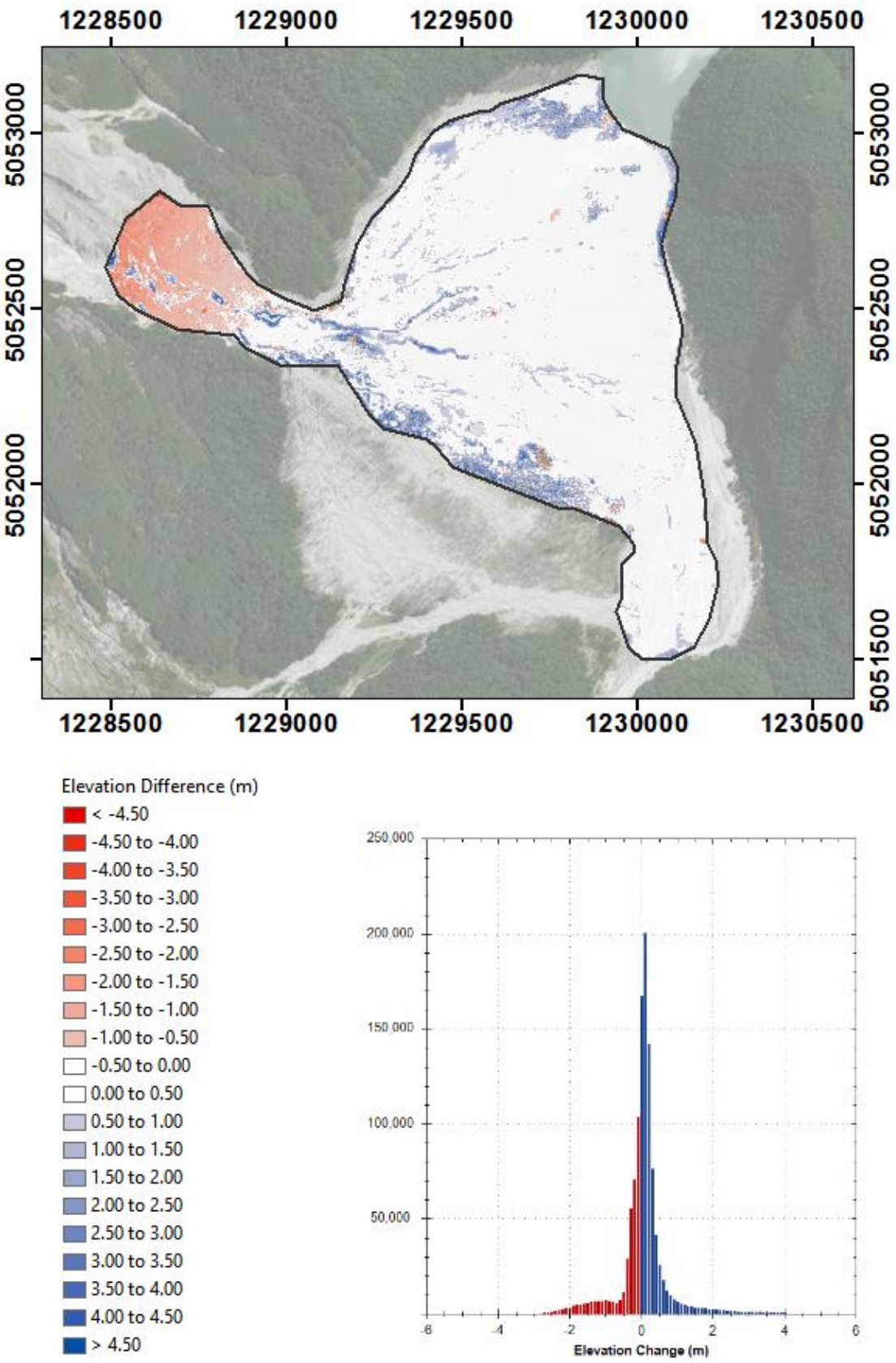

Figure 4.17 TLS-SfM DoD obtained with additional ground control processed within APS, see 4.4.3.5. Whilst error persists in the steep fall section, the toe region is predominantly within 0.5 m. *Projection: NZTM 2000*

4.4.4.1. Standard Approach

A preliminary reconstruction was undertaken (Figure 4.13) utilising a standardised workflow appropriate to the prior information available. In the absence of camera GPS, the disabled alignment setting was used whilst all other parameters selected reflect the approaches outlined in section 3.4.6. Systematic errors are apparent, with significant warping of up to 2 m towards the southern edge of the fan where ground control is limited. In contrast, >85% of the area bounded by the radius of ground control is within 0.5 m. There is noticeable deformation present at the steep source of the fan across all alignments. This area suffered from poor lighting and highly oblique sensor to feature angles, deteriorating the keypoint quality and leading to large deviations of up to 3 m in some locations. The histogram (Figure 4.13) distribution with the beginnings of a bimodal peak, alludes to the presence of the significant errors present at both the edge of the fan and in the steep section.

4.4.4.2. Image Quality

Reconstructions using the edited imagery set show minimal changes against the standard procedure seen in Figure 4.14 across the surface of the fan constrained within the GCP perimeter. There is evidence of potentially reduced doming at the toe, both visually and from the 24% reduction in error at the Checkpoint 2 and 5% at the Checkpoint 1, whilst the severity and spread of areas falling significantly below the TLS control surface is reduced. The secondary peak noted on the histogram in Figure 4.14 is reduced in amplitude, however, the tail extends to a greater magnitude. The reduction in systematic bias coincides with a 3.4% reduction in change to the K1 variable during optimization. This reflects the removal of poor quality images that distort the bundle adjustment, leading to a deterioration in estimated scene structure.

However, the total area within a range of +-0.5 m is reduced when compared to the standard alignment (72.27 % – 71.50 %), contrary to expectations. This potential reflects poor image coverage around the perimeter and centre, where further removal of images degrades the ultimate solution. This can be seen by the isolated halos of >2 m error located throughout that represent the perimeter of regions that could not be effectively reconstructed. This is reflected in an increased SDE in comparison to any other construction of the fan of 0.924 m. The side effect of reducing imagery highlights the

conundrum of whether to include imagery that incorporate a significant percentage of poor tiepoints if these nonetheless introduce new information in regions of sparse data.

4.4.4.3. Pre-Calibration

The GCP placement impacts upon the degree and location of model distortion, which is linked to inaccurate calibration (James and Robson, 2014). To reduce this distortion, a series of iterative reconstructions (0, 1, 2, 3) were created, based upon using an initial camera calibration derived from a previously geo-referenced and optimized scan.

Using the original imagery set resulted in significant reconstruction splicing and so the results for the Image Quality set (n=357) have been used. This image set provides a consistent geometry with no splicing evident until the 4th calibration iteration. The results of the optimum reconstruction (the 2nd iteration) are shown in Figure 4.15. This fails to remove the systematic error but can nonetheless be seen to reduce it, at both the toe and within the GCP radius itself, with 74.62 % of the total area <0.5 m error. LOOCV error remains across the original GCPs, however, the errors at Checkpoint 1 and 2 are reduced by 50% and 25%, respectively, with a reduction in MAE. The K1 parameter undergoes a 0.12% change during the optimization phase of the iteration presented here, in contrast to 1.71% during the High alignment. The impact this progressive calibration approach has upon the interior camera orientation can be tracked by the changes to the radial distortion parameters K1, Figure 4.18.

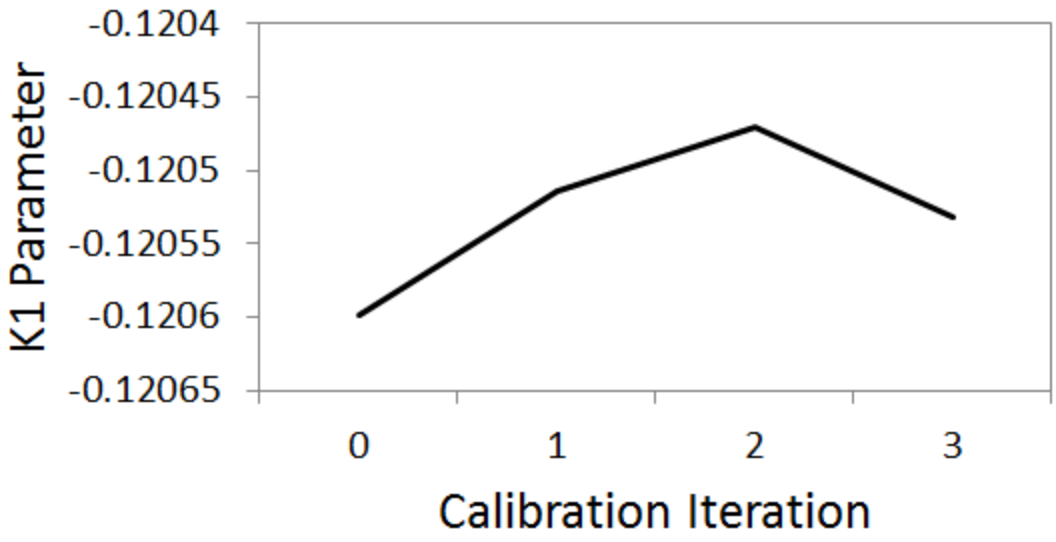

Figure 4.18 Progressive change in the first radial distortion parameter, k1, across the iterative manual calibration process. The 0 calibration represents the initial reconstruction values.

Improvements in the DEM from the initial reconstruction to the second iteration reduce the degree to which the toe of the fan domes. However, upon the 3rd iteration results deteriorate and this reflects a change of direction in optimized value of the radial distortion parameter. Overall results are improved, specifically at the edges of the fan, and it would appear that providing an initial calibration estimate may also act to constrain the bundle adjustment leading to localised errors in areas of poor coverage.

The implementation of this approach is dependent on the identification of a 'good' result during the initial reconstruction, as error propagates through each iteration. This approach appears to refine a theoretically incorrect result as opposed to help determine new local minima. The additional constraints introduced in relation to the image set alterations appear to negate the impact of splicing and would suggest that this approach may be unsatisfactory for implementation over larger scales.

4.4.4.4. Inferred Camera Pose

Figure 4.16 presents results for the use of estimated camera positional data which present an improvement on the standard procedure. An additional reconstruction which introduced a random 3D camera error between 0 – 10 m resulted in an equivalent surface with comparable LOOCV values, with a slight improvement from 0.303 m vs. 0.272 m.

Scans conducted with a pre-referenced camera show a reduction in error at the independent control of 34% (Checkpoint 2) and 7% (Checkpoint 1) when compared to the standard procedure; coupled with a 7% reduction in change required to the K1 parameter during optimization. Additionally, utilising camera GPS as opposed to solving the bundle adjustment with no a priori positional data offers significant time gains, in the order of 35 minutes vs. 355 minutes for a high alignment. This accounts for an 83% reduction in time for a subset of ~ 500 images. The non-linear nature of processing time with increasing numbers would represent a significant time saving over the catchment wide survey. Perhaps most importantly, is that the use of exterior pose information does not result in a reduction in model quality and validates the use, as seen in section 4.2 across the length of the scan.

An additional reconstruction was undertaken using both inferred camera positions and the 2nd calibration iteration as seen in section 4.4.3.3. While in theory, the use of both interior and exterior constraints offers the chance to successfully reconstruct the model with minimal systematic error, in reality the high degree of error in both estimates results in an imperfect solution that fails to build.

4.4.4.5. Extra GCPs

To assess the impact of GCP placement and in an attempt to better represent the fan, the two further manually identified checkpoints were used as GCPs. This approach appears to mitigate the systematic doming, with an increased area of the model falling within <0.5 m to the reference and 77.40 % of the total fan area within 0.5 m. Standard Deviation of Errors is reduced by 10% to 0.792 m. While a LOOCV analysis suggests an increase in the observed error (0.501 m), this effect is masked by the inclusion of additional GCPs which are less accurate than the existing RTK measurements, propagating additional error into the reconstruction (Clapuyt et al., 2016). Despite this, it would appear that low quality control may be used (as in section 4.2) to improve reconstructions in locations where 'some' a priori information, is better than none.

Figure 4.17, reveals the characteristic doming effect in the sparse cloud that appears to arise from compensation in the estimated interior camera geometry. However, there is a slight banding of increased error in the region of Checkpoint 1, which is reflected in the leave-one-out analysis. This systematic error is to be expected, as improved distribution of control fails to eliminate the sources of systematic error. Checkpoint 1 experiences significantly larger error than that of Checkpoint 2, despite closer proximity to the original GCPs which is potentially indicative of poor distribution of GCPs.

While the extra GCPs provide increased horizontal coverage of the fan, there remains a considerable lack of control in terms of in the z dimension. Errors in the steep section of the model (i.e., the landslide gully and adjacent hillslopes) are likely to relate to poor image quality (due to shading) and the highly oblique surface angles. The degree to which systematic error appears to have been reduced through the addition of further control suggests that the lack of vertically variable control is the ultimate limiting factor.

A final consideration is highlighted by the mean error, which under both this scenario and that of the pre-calibrated reconstruction, result in an elevated value indicating an overall bias. In reality, the reconstruction is positive trending due to the inconsistencies in the ability of TLS and SfM to reconstruct vegetation and the major gullies. The seemingly low mean values presented for other reconstructions are likely masking this with the significant regions of systematic error that cancel out these anomalies.

4.4.5. Summary of experimental design variations

The experiments relating to tiepoint quality, interior and exterior constraints reveal varying degrees of influence upon the quality of results. The impact of problem imagery remains difficult to quantify, in part due to complicated impacts of introducing both spurious tiepoints and new informative tiepoints, pertinent in low density capture regions. Most successful is the introduction of a priori interior calibration estimates, which reduce checkpoint errors and increase the percentage of the model falling within 0.5 m of the reference data. The most practical benefit, however, is felt through the introduction of exterior pose estimates, which dramatically reduce the computational burden, of specific importance to the Dart survey. The combination of exterior and interior pose estimates together, unfortunately, fails to produce viable reconstructions, perhaps unsurprising, when constraining the bundle adjustment by introducing two fundamentally incorrect conditions.

4.5. Discussion

4.5.1. Systematic error present within reconstructions

Systematic errors are evident within the 2014 catchment wide scan, easily identified through ICP analysis and comparisons to existing LiDAR data. Systematic error prevalent within the literature (James and Robson, 2012; Westoby et al., 2012; Fonstad et al., 2013) is especially evident in regions outside of ground control. This can be seen to manifest as a result of inaccurate initial sparse clouds that require extensive correction during the optimization process. Alongside this, the consistency of camera locations is found to impact as dramatically inaccurate extrinsic parameters can lead to the identification of incorrect local minima during the bundle adjustment. While such an approach provides an insight into the SfM pipeline, the lack of spatial control limits the

results to point recording based inferences when considering the 2014 catchment wide data, addressed by the experiments conducted on the fan.

There is no clear systematic pattern of error observed through the ICP analysis of the catchment, however, the optimization phase indicates the presence of a wave like pattern. The nature of this correction deviates from the doming pattern observed across multiple studies performed at smaller scales in more appropriately shaped blocks. To determine whether optimization results were a result of the survey strategy or likely an indication of the challenges faced conducting a corridor survey, the systematic error removed during the optimization phase of the debris fan was assessed (Figure 4.19).

The standard reconstruction exhibits a concentric ring pattern, bounded by the perimeter of the seven initial GCPs, past which the correction is inferred as the rate of correction applied per distance from radial centre accelerates. This is representative of the "doming effect" and specifically the impact of varying the K1 radial distortion parameter as seen previously by James and Robson (2014). The corrections applied in the extremities are non-uniform and significant at >2 m.

With respect to additional GCPs, the radial centre has shifted along with the degree of correction required. The improved control leads to a more accurate initial linear transformation, which requires subsequently less correction during the optimization phase, as seen from the 25% gradient reduction in the radial plot. There is also a greater distribution of both negative and positive change which reflects the improved effective real surface intercept of the SfM construct. Ultimately, while increasing the extent of the GCPs improves results, the doming effect remains, which is an issue when the distribution is limited by control. Results reinforce minimising systematic error is dependent upon the accuracy of the initial sparse alignment. A perfectly calibrated (intrinsic and extrinsic) camera would need only an affine transformation to be projected in a real world co-ordinate system and thus any changes before and after the optimization process implemented in APS would be minimal, reducing the reliance on an extensive GCP network and increasing applicability to hostile sites in a corridor setting. This is potentially achieved through the adoption of convergent imagery (Wackrow and Chandler, 2011; James and Robson, 2014; Rupnick et al., 2015; Woodget et al., 2015; Dietrich, 2016); in such an instance the optimized DOD would be expected

to produce a primarily flat surface, with significantly reduced warping at the radial edges.

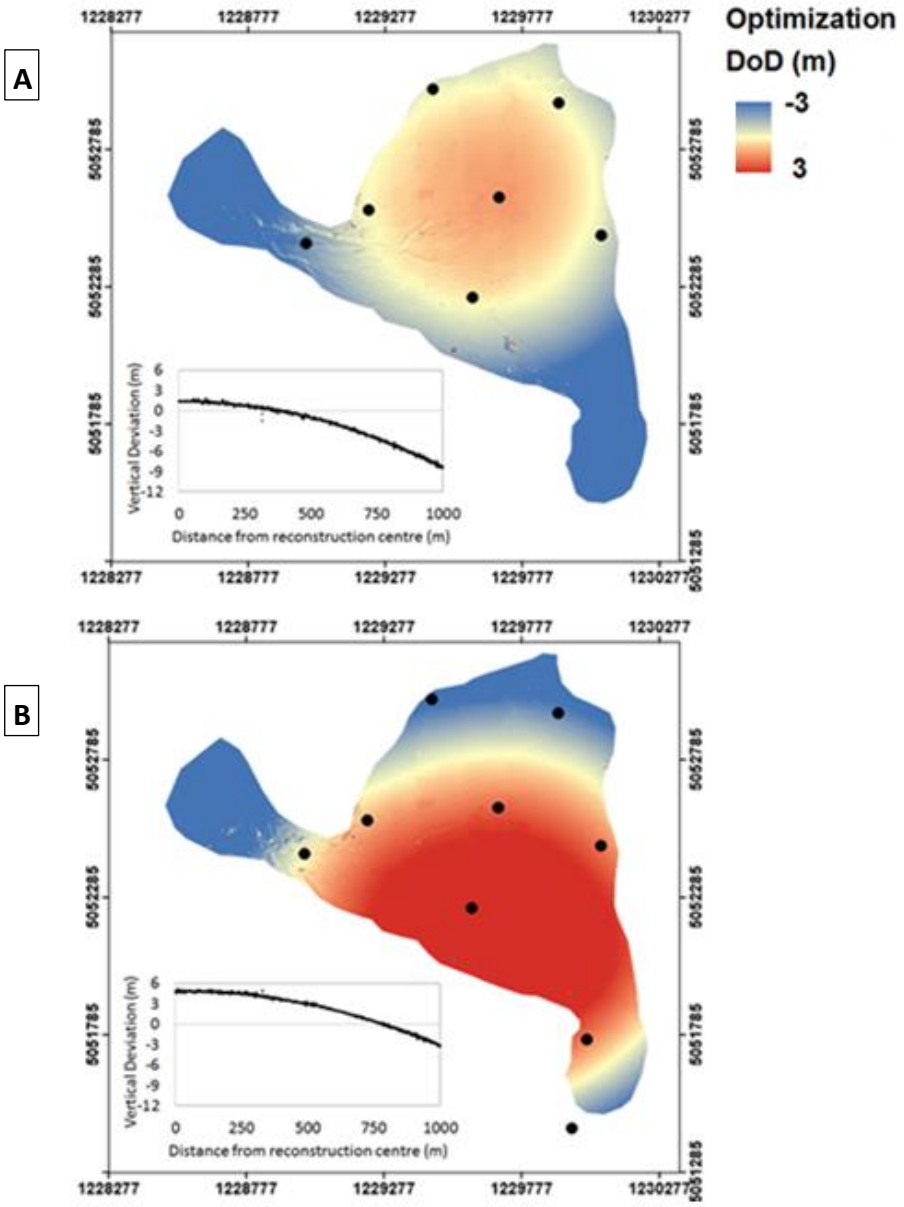

Figure 4.19 Dem-of-Differences derived pre and post optimization within APS to reflect the impact of the lens correct upon the initial sparse cloud geometry and equivalent Lens correction (m) plotted against radial distance (m) from centre point of correction for a) High Alignment, b) Additional GPS. The DoDs observed (figure 4.13 and figure 4.17) suggest the elimination of the doming effect, the optimization step shows otherwise. *Projection: NZTM 2000*

Finally, it is important to consider that the work undertaken on the fan is subject to localised systematic error, resulting from the distribution of the control used and random errors, relating to areas of shading and high contrast. In contrast, the degree of systematic and random errors over the entire catchment is influenced by the quality of imagery and distribution of control across the entire reach. The impact of these

differences may be observed through comparing both a local and regional model of the fan to identify variations, see Figure 4.20.

The overall geometry of the fan appears to exhibit rather less geometric variation when compared to the reference TLS dataset, due to an improved distribution of control. However, the accuracy of GCPs within a radius of the fan centred reconstructions is, by comparison, degraded. This is reflected by comparing the two SfM datasets whereby the systematic error at the toe of the fan is clearly defined, however there is an increase in noise when analysing the area within the GCP perimeter for the fan. In the context of this study, this perceived increase in systematic quality vs. a decrease in random quality is justified due to the degree of errors introduced; local surface is reduced in quality by 10-20 cm, vs. regional surface improvements of up to 2 m.

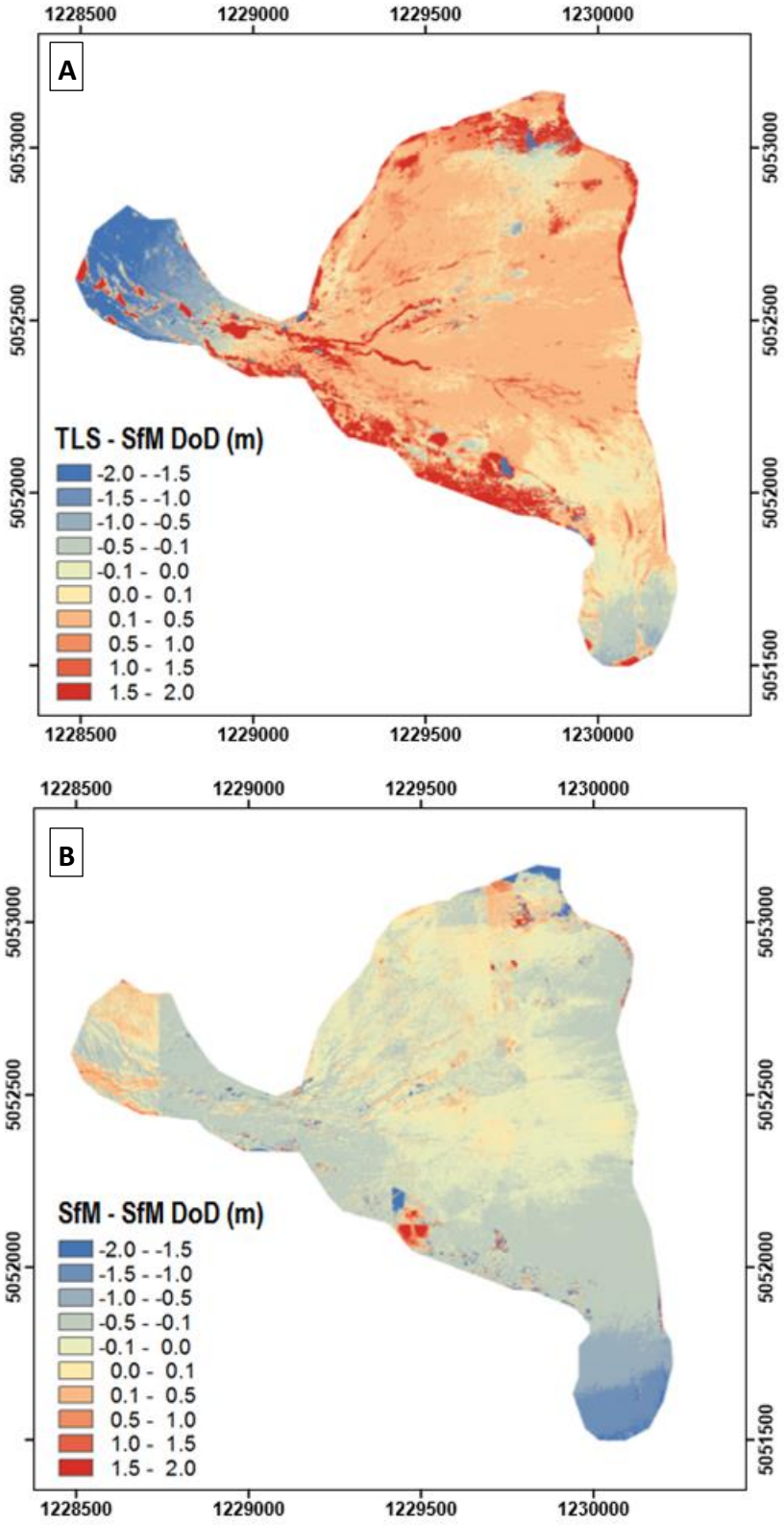

Figure 4.20 A) TLS-SfM DoD and B) SfM-SfM DoD, comparing the fan study site topography obtained from the catchment wide scan against the control data and the initial High Alignment. The scale of reconstruction is seen to result in differences in whether the area is systematically optimal, or locally optimal. *Projection: NZTM 2000*

4.5.2. Localized error present within reconstructions

In line with previous studies, SfM reconstructions struggle to accurately capture areas of particularly steep and poorly lit terrain, along with vegetated regions and water bodies (Niethammer et al., 2012; Luciier et al., 2013; Gomez-Guitierrez et al., 2014b). Minimizing these local errors in an attempt to improve tiepoint quality has been undertaken upon the study section of the debris fan. To reduce subjectivity in the process, only images of over 90% vegetated or wetted area were removed from the analysis. In reality, this threshold must be set lower, with image inclusion only occurring in instances where the remaining scene provides new information, as determined by the user. The steep fan section is associated with errors of >3 m, both positive and minus, suggesting significant local variability, primarily through an inability to detect keypoints in the heavily shaded region. However, the impact of a lack of significant control in the z dimension must be considered, especially for regions that are at least 200 m above the effective GCP plane. Particularly steep sections are captured at relatively highly oblique feature to sensor angles, typically most evident in the radial edge of photographs as they remain un-viewable from nadir sensors when directly above. These prevalent errors are mimicked in the larger gullies and channels along the surface of the fan, once again impacted by poor lighting and steep slopes. Future studies may alleviate such error through capture under diffuse lighting conditions and obtaining additional oblique angle images to provide relative nadir capture of the slope planar surface.

Significant errors were evident in the form of a series of holes in the data. These are likely to reflect poor flight plan management and a low image density. In the absence of sufficient overlap, particular regions of the model cannot be reconstructed due to the failure of correspondence matches, leading to areas recorded as >2 m error. This effect was seen not only on the debris fan, but also when constructing the gorge as part of the entire catchment. With the use of UASs, this effect may be avoided through the use of flight planning software (Turner et al., 2012; Lucieer et al., 2014a). At the catchment scale when using a helicopter navigated via a series of waypoints on a tablet screen renders stringent flight planning unfeasible. In line with the literature (Westoby et al., 2012; Fonstad et al., 2013), overshooting is advised, necessitating post capture image removal to reduce computational time, as opposed to leaving data holes. Overshooting may be done through increasing the image capture frequency or increasing the number

of flight paths; although any such approach should be treated cautiously to maintain ideal camera to feature base line distances (Marcis, 2013). The impact of overshooting over such a large area does, however, raise issues relating to flight height/ speed, image format and the computational resources available.

4.6. Conclusions

This chapter has produced digital terrain products for a 35 km reach of the Dart catchment utilising a SfM survey captured from a helicopter. In so doing, there are a number of key findings to emerge:

1) A nadir orientated survey is able to capture the topography of a river catchment of length 35 km and area 44 km^2, with ICP MAE errors of the order of 0.3 m. There is a strong caveat here, however, in that the data are only reliable within the radius of ground control, with significant deterioration outside this perimeter. It is established that future work would benefit from identifying ways to minimize the reliance on ground control.
2) The quality of tiepoints, in terms of fidelity and scale is noted throughout these experiments. The High alignment settings are recommended to enable keypoint extraction, without degrading accuracy. Low alignment settings remain a viable solution when constrained by GCPs, especially for larger datasets, however they must be closely monitored to assess both the increased local surface variability and the impact that varying feature scales may have across varying landscapes. The removal of problematic images has been shown to partially reduce doming in the southern toe of the fan reconstruction, however this has introduced gaps in the data given the low percentage overlap in this image set. It is important to consider the trade-off between improved model fidelity and reduction in computational times against a reduction in coverage and the resulting holes in the data.
3) Introducing interior constraints into the bundle adjustment, through a progressive pre-calibration improves results, reducing the correction required during the optimization process. The calibration process has been shown to reduce deformation by 50% at a distance of 500 m and 25% at 800 m. Ultimately

this process acts largely to improve results outside the perimeter of the bounding GCPs when no further information is available. The method is unable to eliminate the doming effect fully from the sparse cloud and is surpassed in terms of accuracy by any reconstruction with a more complete GCP distribution. When performing a progressive initial calibration the K1 parameter should be tracked to identify the point at which results begin to deteriorate. It must be acknowledged that such an approach to pre-calibration merely acts to refine results within incorrectly identified local minima and will never mimic the accuracy of traditional metric cameras. Furthermore, within the confines of the available APS version used here, the use of prior interior constraints fails to successfully reconstruct when combined with pose estimations. In terms of pure practicality, the use of the Reference setting within APS and hence providing exterior pose estimates is a necessity to enable the use of High alignment and successful reconstruction within CPU limitations.

4) Camera GPS is highly recommended, with no evident degradation of model accuracy at a fraction of the processing time and is robust to errors tested up to 10 m. The processing time without initial pose information is based upon a factorial relationship, assuming each image must attempt to match with each other at least once, leading to a non-linear increase in time relating to image numbers. Considering the computational requirements for larger data sets, camera GPS provides an opportunity to reduce processing times, or perhaps more importantly considering the inaccuracies of the low alignment setting, allow for higher alignment accuracies to be used where previously unrealistic. The combination of iterative calibration and estimated camera GPS is prone to splicing and therefore future work is recommended to focus upon using camera GPS only as a result of the significant time benefits.

At the scale of the debris fan Slipstream, the SfM pipeline has proved successful in capturing the topography down to decimetre scale with comparable results to that of a TLS survey. Results across the scans are of the order of ~72% <0.5 m error, despite the impact of the steep fan source and systematic toe deformation. A ground pixel resolution of 6 cm is deemed sufficient to detect keypoints and capture topographic features relevant to the interests of the study. Whilst a series of refinements has been

suggested, targeting the key stages of the pipeline, improvements are proven, however, to be of an order of magnitude below the advantages achieved from adequate GCP placement and improved camera network geometry.

4.7. Future Work

This chapter has described and tested a workflow for aerial SfM photogrammetry as a means for capturing topography in remote and wild terrain. Whilst ground control remains a necessity, future work upon the Dart should first be aimed at eliminating systematic bias from the initial sparse cloud and so reducing the reliance upon this. Major enhancements in this direction have been made using convergent image networks (Rupnick et al., 2015; Woodget et al., 2015; Dietrich, 2016). This may be tested in a controlled environment to determine the optimum camera geometry for future surveys. Additionally, subsequent surveys will adopt the use of camera GPS to take advantage of increased computational efficiency.

5. Chapter 5: Minimizing systematic errors within the Structure-from-Motion pipeline: The use of flume based, scaled experiments to determine the optimum camera network geometry

Abstract

Chapter 4 highlighted the presence of systematic errors within Structure-from-Motion derived reconstructions of the 2014 Dart River survey. Results obtained elsewhere, suggest that these effects are likely to relate to inaccurate camera calibration (interior and exterior) compounded by sub-optimal distribution of external constraints, i.e., the GCP targets. The nature of the deformation and the magnitude of errors observed (particularly beyond the extent of ground control) suggest that a desirable goal is to minimize the reliance on ground control as an a posteriori solution to remove such systematic errors. Instead, the focus should rest on supporting a more robust solution to the initial bundle adjustment and hence improve the reconstructed structure prior to subsequent camera recalibration using exterior constraints.

With this in mind, a series of scaled experiments using a cost-effective physical model were used to explore the optimal camera network design for future surveys. These experiments focused in particular, on the benefits of combining oblique and vertical photography in a convergent camera network. It has previously been observed that convergent camera networks improve depth estimates and geometric accuracy during the bundle adjustment, and mitigate systematic distortion in the resulting point cloud (Wackrow et al., 2011; James and Robson, 2014).

Here, a series of scaled experiments was used to evaluate the optimum camera angles for the reconstruction of prototype braided river morphologies. The use of multiple camera angles does, however, add a severe computation overhead. As such, the effects of additional survey parameters, in particular, image density and image format were also examined to see if an effective trade-off could be achieved. This focuses on how the benefits of additional imagery can be offset with lower quality image products and/or lower density coverage.

The results presented indicate that convergent photography is an effective means to mitigate bowl/doming type systematic errors in the initial bundle adjustment and in turn relax the reliance on posterior camera calibration. Assessment of reconstructions of flume channel topography suggest that check point errors can be reduced by 40% (0.010 m to 0.006 m) using convergent camera networks angled at 30^0. Beyond this angle, however, the addition of severely oblique images was found to introduce unwanted noise.

It is proposed that this survey design can be best realized by employing two angled camera on a single rig, rather than flying opposed flight lines. This does, however, effectively double the image count creating a major increase in the computational burden, as there is a direct correlation between image numbers and processing time during the bundle adjustment. Simulated thinning of the image network identified that optimal model reconstructions can be obtained with as little as 69% overlap using convergent imaging networks, although initial overshooting is recommended as it is easier to remove images then re-obtain images after the survey. The use of compressed JPEG imagery was also found to provide an adequate compromise between computational time (reducing reconstruction times by 19% cf TIFF format imagery) with only minimal changes in model accuracy (MAE OF 0.0008 m across all scans). Results presented demonstrate that approximately equivalent quality surface reconstructions of the flume topography can be obtained using only 16 JPEG images as opposed to 48 TIFF images, providing a 73% reduction in processing time.

The results of the laboratory simulations are finally extended by considering aerial reconstructions of the township of Glenorchy, for which there is abundant ground control and an independent LiDAR survey for model assessment. Comparable tests to those undertaken in the flume are presented to demonstrate the scalability of experiments and successful implementation at the field setting, when considering the impact of imaging networks and GCP distribution restrictions. Comparisons between sparse cloud reconstructions using convergent and vertical image networks demonstrate significant improvements in the identified scene structure, with check point errors reduced by over 50% (down to 0.177 m from 0.404 m). The difference

between the camera networks is revealed by the interaction between the exterior pose estimates and the compensating focal length parameter.

5.1. Chapter Objectives

5.1.1. Identification of operational concerns

Accurate topography derived from Structure-from-motion surveys requires optimal distribution and density of ground control across the area of interest. This is particularly important, as peripheral areas of the reconstructed structure have been found to be prone to systematic distortions, resulting from the miscalibration of intrinsic and extrinsic camera parameters (Wackrow and Chandler, 2008; Wackrow et al., 2011; James and Robson, 2014). Such systematic errors have been linked to poor estimation of the radial distortion and focal length parameters. This stems from compensation between the intrinsic and extrinsic camera parameters in the SfM bundle adjustment which is solved initially using inner constrains only (James and Robson, 2014; Eltner and Schneider, 2015; Clapuyt et al., 2016). While it is possible to mitigate some of these effects post-hoc, through a secondary calibration of the intrinsic camera parameters using ground control constraints, improvements in the reconstructed topography will be biased to the distribution of external constraints used with further distortion likely in areas of sparse ground control. Understanding how to reduce the presence of such distortion is therefore vital to our ability to successfully translate the survey strategies employed in intensive but, localized UAS studies, to the catchment-scale corridor mapping exercises applied here with their much lower densities of GCPs and highly constrained deployments.

Convergent camera networks, comprising either oblique-oblique or oblique-nadir imagery have been suggested to reduce the impact of systematic distortion through improved geometric accuracy during the initial bundle adjustment (Wackrow and Chandler, 2008; Wackrow et al., 2011; Marcis, 2013; James and Robson, 2014; Woodget et al., 2014; Dietrich, 2016). Such convergent photogrammetric networks are less sensitive to radial distortion when observing the same target from paired viewpoints. With vertical imagery, these correction errors lead to inaccurate depth perceptions, which manifest themselves as a ‘dome’ or ‘bowl-shaped’ systematic distortions in the sparse point cloud (Wackrow et al., 2011). Within APS a non-linear block transformation (optimization) may be applied, a posteriori to re-estimate the lens distortion using additional constraints in the form of known GCP targets. This approach is effective if abundant and well-distributed ground control is available. However, where ground

control is sparsely distributed, there is the potential for significant local variation in the quality of the reconstructed surface, as uncertainties in the bundle adjustment manifest in complex distortion patterns. As such, good practice should focus rather less on post-hoc solutions and instead on strategies to improve the accuracy and robustness of the initial bundle adjustment (Cornelis et al., 2002).

The optimal design of convergent camera networks for SfM reconstructions has so far received relatively little attention. Studies using traditional stereo photogrammetry, such as Karara and Abdel-Aziz (1974) initially proposed minimising camera convergence, while Wackrow and Chandler (2008) found optimal results using two cameras angled inwards at 15^{o}, creating a relative imaging geometry of 30^{o}. More recently, Marcis (2013) compared singular multi-angle sensors of a planar surface and found optimal results normal to the plane and inclined at 65^{o} while model quality was observed to decrease as viewing angles were lowered below 45^{o}. There is therefore an urgent need to evaluate convergent camera geometry for SfM studies and this is likely to be particularly important for studies such as here, where post-hoc camera calibration is based on only sparse ground control.

5.1.2. Research Design

Standard approaches to establishing a convergent network employ the use of multiple shallow oblique angles, often shot by a single camera. To operationalize this here, a multiple sensor system, based on a nadir-aft-nadir geometry similar to that used in satellite stereo-imaging (e.g., the Worldview satellite series) is proposed. This selection reflects a conscious decision to maximise the imagery obtained during the aerial survey, which constitutes a major financial outlay for the project. Acquiring imagery from two perspectives simultaneously may result in a degree of overshooting, but this can be managed by later filtering the image set if necessary. The use of additional sensors does, nonetheless, introduce a number of secondary concerns, the most obvious of which is the doubling of image numbers which will pose computational overheads during both the bundle adjustment and depth mapping.

Designing the optimal camera network would ideally be evaluated using multiple aerial surveys that reflect the scale and geometry of the scene to be captured. Such a strategy is prohibitively costly, so here an alternative approach is used based on scaled, flume

based experiments. This provides a cost-effective and flexible means to test hypotheses without recourse to repeat helicopter flights under field configurations. While as indicated in Chapter 5, there are a number of options to refine the survey approach used in 2014, many of these avenues are difficult to control in the remote and hostile terrain of the study area, i.e., improving lighting conditions and the GCP density and distribution. Refining the camera network geometry in order to provide a more robust bundle adjustment does therefore represent an appropriate first step, which may relax the need for additional exterior constraints to calibrate the intrinsic cameras post-hoc.

The research reported in this chapter is therefore aimed at reducing the systematic distortion apparent in the 2014 reconstruction. This will be achieved through a number of experiments aimed at determining the optimal image capture methodology, focusing on imaging angles, capture frequency and digital image format and evaluated in respect of post-processing time, logistical constraints as well as reconstruction quality. These goals are met through addressing the following research objectives:

- **Objective 1**: Identify the optimal camera network configuration to reduce systematic error and relax reliance upon ground control.
- **Objective 2**: Identify ways in which the workflow may be refined to improve efficiency when applying to a catchment scale project.
- **Objective 3**: Validate the use of a flume and scaled experiments as a cost effective means to inform future survey strategies.

These objectives will in turn be met through a series of experiments which are outlined in Table 5.1 below.

Table 5.1 The three principal experiments undertaken in this chapter, detailing their rational and potential benefits in developing a new strategy for the 2015 Dart survey.

Experiment	Application to SfM	Application to the Dart/ Large Scale
Convergent Imagery	Reduces impact of systematic errors resultant from incorrect camera calibration, particularly in regions beyond the extent of ground control (Wackrow and Chandler, 2008; James and Robson, 2014; Dietrich, 2016).	Remote hostile environment with severe limitations on distribution of ground control to help minimize systematic errors
Reducing Imagery Numbers	Reduces computational burden of the SfM process if an area is kept constant, or provides expansion of surveys through greater flight speeds	Previous survey contained 5002 original images, subsequent dual sensor approach will double photos. The computational impact of additional photos is non-linear within APS. Will allow for greater flight speeds to cover the ~35km length of the Dart
JPEG Compression	Reduces computational burden and allows for faster file write speeds during capture process	Enable a reduction of processing times for large datasets, through increase in pixel size in homogenous areas. Allow for reading/writing speed required during flight with capture rate of 1Hz and GPS.

5.2. Scaled Structure-from-Motion Experiments

The experiments reported here were designed to image a simulated braided river network developed on a 2 m x 1 m indoor flume table. An independent 3D model of this surface for comparison was obtained using a short range TLS survey. Three experiments were undertaken and benchmarked against this TLS dataset as follows aimed at achieving objective 1 and 2:

A. SfM reconstructions based on nadir orientated and multiple convergent image networks using an aft-nadir set up, of secondary angles 15^{o}, 30^{o} and 45^{o} off nadir.
B. Reconstructions generated using systematically thinned image sets to represent different flight speeds/ capture frequencies.

C. Reconstructions generated using both lossless TIFF and compressed JPEG images.

5.2.1. The flume table and simulated topography

The indoor flume provided an ideal test scenario that is both cost effective and offers the ability generate synthetic topography of broadly similar morphology (in terms of topographic amplitude) to the prototype system.

Image capture was facilitated through a metal frame suspended above the flume, from which a directional camera mount could be attached and the sensor orientated (see Figure 5.1). This provided a consistent platform for image capture at multiple locations above the flume and from a wide range of viewing angles.

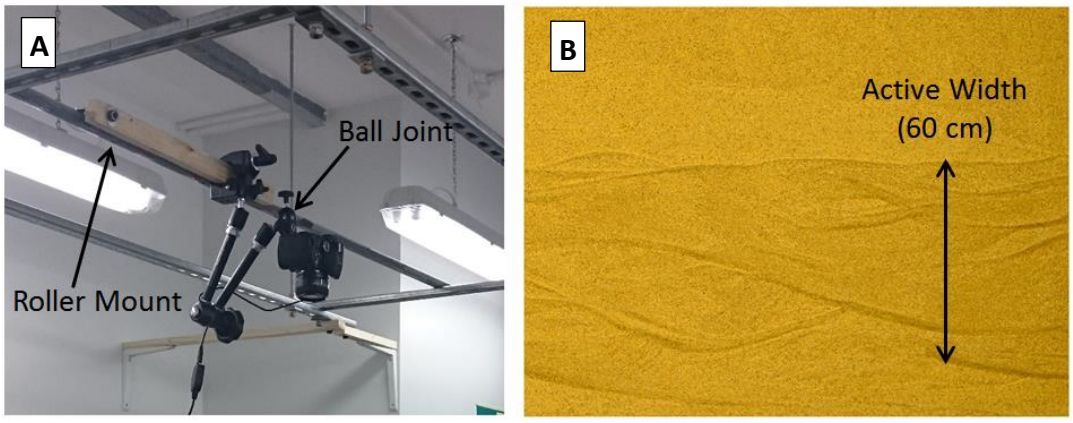

Figure 5.1 A) The camera, mount, roller and ball joint used to capture imagery and B) a corresponding image of the flume from a height of 1.35 m, displaying a braided reach of active width of 60 cm.

The frame mounted above the flume restricted the camera height to just 1.35 m during this survey which compares to a mean Above Ground Height (AGH) of 350 m in the field. This provides an approximate scaling ratio of 1:260, which is used to inform the generation of the braided topography within the flume. The flume used is an Armfield S17 River Flow Simulator that comprises a 2 x 1 m steel tank, with a recirculating (water only) pump capable of providing flow rates of between 1-15 l/min. The tank can be inclined using a simple jacking system. For these experiments, a braided river network was generated in a medium sand (d50 = 1 mm) substrate, under steady flow (8 l/min)

with manually recirculated sediment and a gradient of 0.05. The topography evolved from an initial plane bed channel of diameter 0.2 m and was allowed to migrate laterally, to develop a steady-state braided planform after 4 hours. The resulting channel had an active width of typically 0.6 m with a vertical height range laterally of between 0.3 -1 cm. Taking a representative width of the Dart as 1.5 km, this broadly scales geometrically as an equivalent bar top-to-channel bed difference of ~1-3 m, which reflects the typically observed morphology.

The resulting scaling distances between the flume and field can be seen in table 5.2. The flume morphology studied may be best seen to represent a 500 m reach of the Dart. As such a number of caveats are present, primarily, that some of the issues that are specific to corridor mapping are likely to be redundant in the subsequent analysis. This study represents an attempt to minimize systematic errors and a reliance on ground control by assessing a convenient photogrammetric block. Furthermore, it is further difficult to directly scale and compare error values as errors scale non-linearly beyond the extent of ground control.

Table 5.2 Example of scaled distances between the flume and the Dart River based on a scaling ratio of 1:260, which is imposed based upon the maximum camera to flume bed height.

Feature	**Distance in the Flume**	**Distance in field**
Change detection	0.005 m	1.3 m
Change detection	0.01 m	2.6 m
Distance from GCP's	0.1 m	26 m
Distance from GCP's	0.5 m	130 m
Distance from GCP's	1.0 m	260 m
Camera Height	1.35 m	350 m
Image Spacing	0.14 m	40 m
Bank Height	0.003-0.01 m	0.8 – 2.6 m
Active Width	0.6 m	1.5 km
FOV	0.75 x 1.15 m	196 x 294 m
Length of Reach	2 m	500 m

The experiments required data capture over multiple days and given the desire to compare reconstructed topography to that surveyed by TLS, the flume was therefore allowed to dry before imaging. Such an approach limits an assessment of reconstructed wetted surfaces. Diffuse lighting conditions were provided by eliminating external sources of light to the room and using overhead indoor illumination. Despite this, shadows were present due to cut banks.

5.2.2. SfM reconstruction of the flume

A Nikon D3500 DSLR camera mounted via a ball joint to the frame roller was suspended above the flume (Figure 5.1). The camera angle was set with the use of a spirit level and protractor using a base plate of the camera attachment as a reference guide. While this approach to orientation is somewhat approximate, it was deemed representative of the variation likely in the aerial survey given the unstable attitude of the helicopter and viewing geometry. The horizontal spacing between image capture was initially set to 0.14 m, scaling relative to the distance covered by a helicopter travelling at 90 knots and a capture imagery of frequency of 1 Hz (40 m). To achieve this spacing across the area of interest, the camera mount is fixed at three locations above the flume and then rotated. In so doing, a camera network based on four transects was obtained as shown in Figure 5.4. The footprint of the vertical photography was 0.75 x 1.15 m, which therefore scales to 196 m x 294 m, and is comparable to that of the 2014 survey (200 x 300 m).

Imagery was captured in a RAW NEF format and subsequently converted into a lossless TIFF format to enable use within APS. The photogrammetric reconstructions undertaken using APS, were performed using the 'Disabled' alignment setting (see section 3.4). This setting introduces no prior information of image pairs into the bundle adjustment and is used in the absence of camera co-ordinates. While appropriate for the small camera networks used in these experiments, this setting creates an unreasonably high computational burden for problems involving a large number of images such as the field surveys.

The sparse clouds generated were geo-referenced using a network of control points (see section 5.2.4. below) and then optimized to minimize lens distortion. Subsequently, reconstructions were densified via the pairwise SGM algorithm discussed in Chapter 3,

which provides pixel-wise matching at comparable resolution to a TLS survey. The resulting sparse clouds comprise of ~1.13 x 10^4 points while the dense clouds are typically of the order of 3.7 x 10^6 points and as such provide more detailed models of the channels. Despite the relative homogeneity of the surface, the SfM reconstructions are able to successfully reproduce the flume surface with no obvious gaps in the data model (Figure 5.2).

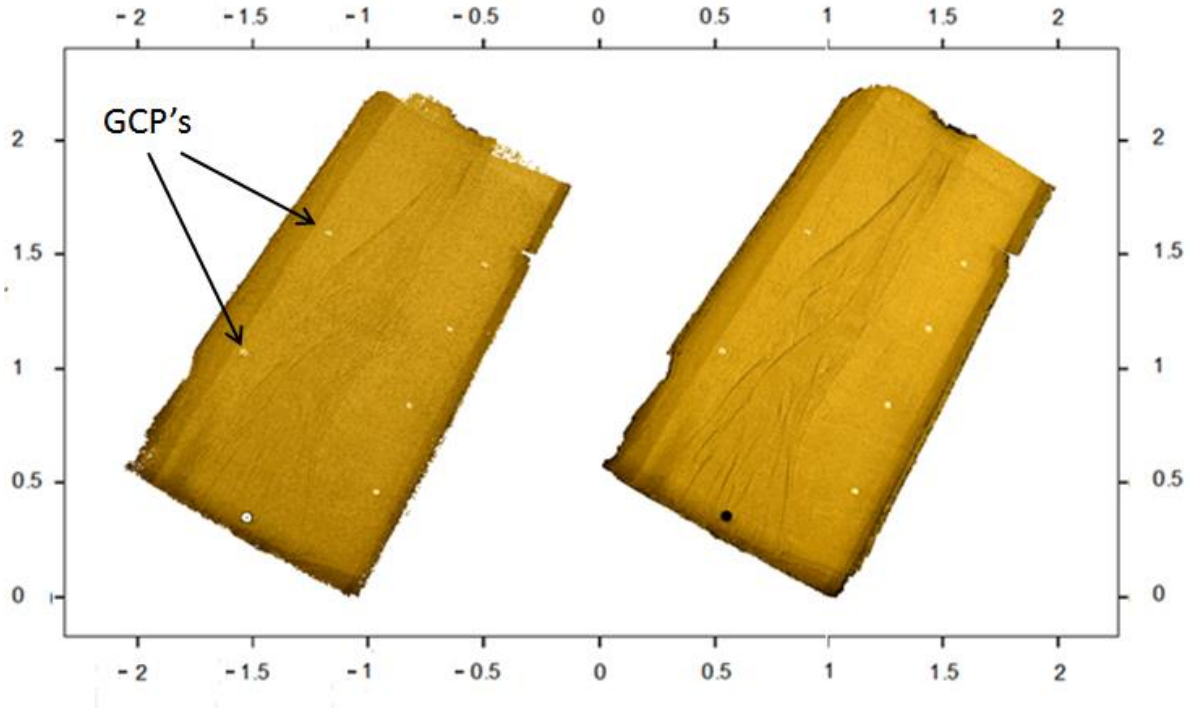

Figure 5.2 Sparse and Dense reconstruction of the flume within APS, consisting of 7.74 x 10^5 and 1.39 x 10^7 points respectively.

Areas of shading are apparent on the banks of the more prominent channels, highlighted in Figure 5.2 by the regions of dramatic contrast. This is significant as SfM uses visible light to perform keypoint identification and feature matching, resulting in reduced surface fidelity. Point clouds were exported in a flat text (x,y,z,r,g,b) format (.pts) and then processed using the Topographic Point Cloud Analysis Toolkit, ToPCAT (Brasington et al., 2012), to generate a 0.002 m resolution decimated point cloud based on the local minimum elevation.

5.2.3. TLS surveying of the flume

The TLS survey employed a Leica HDS6200 scanner. This is a continuous wave scanner with the following specifications: wavelength of 910 nm; an operational range of 0.4-79 m; angular resolution of 125 μrad; 360^0 x 270^0 field of view; and a maximum scan rate of over 1,00,000 points per second. A total of three scans were conducted at approximately 120^o viewing angles to provide overlapping 3D coverage. The scanner resolution was set to 'ultrahigh' providing a point spacing of 1.6 mm at 10 m. Each scan took ~20 mins to acquire and generated a point cloud comprising ~2.5 million observations. The TLS scans were geo-referenced, cleaned and co-registered using Leica Cyclone using a global rigid body transformation based on a minimum of three targets distributed throughout the laboratory. The registered point cloud was then exported and decimated using ToPCAT to generate an equivalent 0.002 m model as derived using SfM.

5.2.4. The Control Network

A consistent coordinate system is required to enable comparison between the SfM and TLS reconstructions. Here, a relative Cartesian system was adopted, set out using a Leica Flexline TS06 total station (accuracy of 2 mm + 2ppm) based on an arbitrary azimuth and datum. A total of eight Ground Control Points were surveyed relative to this coordinate system with four (TL2, TL3, TL4 and TR2) used to register both the SfM and TLS datasets, see the primary GCP distribution in Figure 5.4. A set of coins, with a definitive cross hair observable from photographs and the total station were used for this purpose (Figure 5.3).

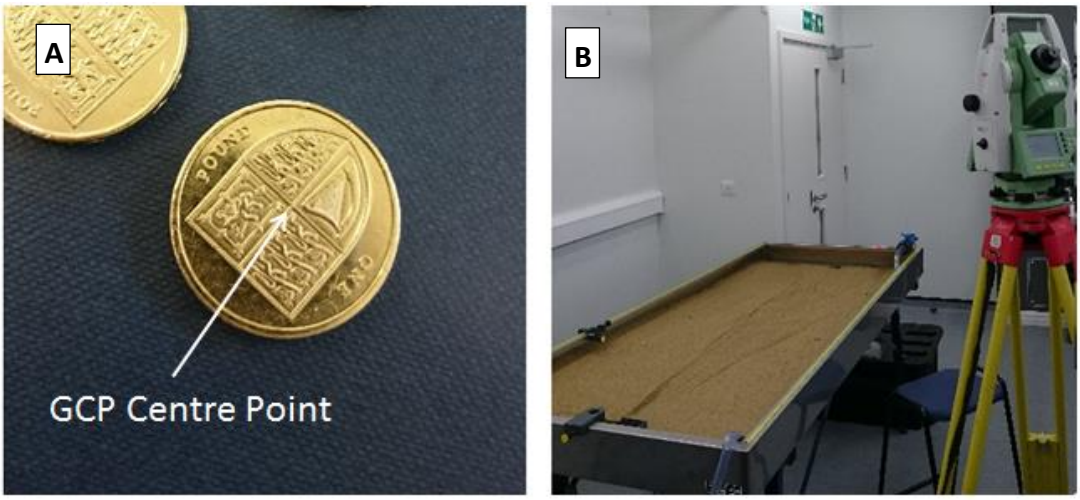

Figure 5.3 A) The GCPs used within the flume within this study. The cross hair of the shield was identifiable from the Total station as well as within imagery captured of the flume B) Total station used to provide referencing of the GCPs within the flume.

The GCPs were located across the flume but beyond the scope of the active channel width. The GCPs listed beyond the extent of the primary GCP distribution serve two purposes; (1) to act as ICPs for independent error analysis and; (2) to provide secondary GCP distributions that are aimed at exacerbating the systematic error (Figure 5.4).

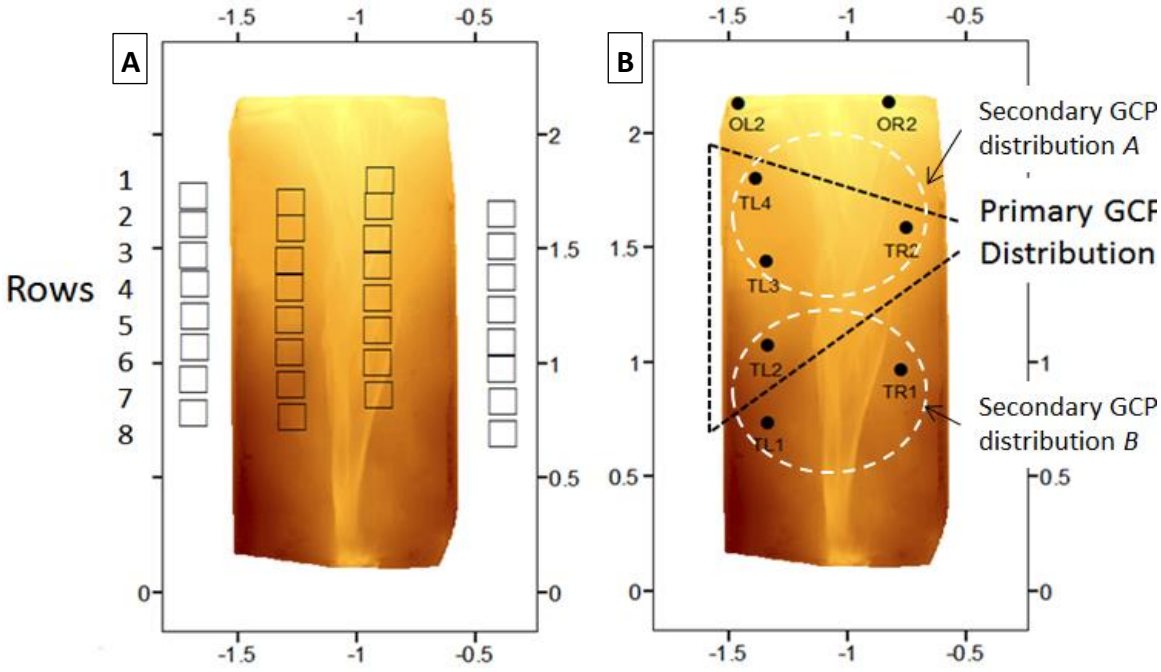

Figure 5.4 A) Plan view of the flume DEM displaying image capture locations. Rows 1 and 4 were used to capture both nadir and oblique imagery. B) GCP and ICP locations within the flume, with the primary GCP distribution used unless specified otherwise.

Two secondary GCP distribution configurations are selected, distribution *A*, consisting of GCPs TL3, TL4 and TR2 and distribution *B*, consisting of TL1, TL2 and TR1. These provide independent control at increased distances to the GCP perimeter in comparison to the primary distribution. For example, the distance between the primary GCP distribution perimeter to point OR2 is in the order of 0.25 m. This distance is increased to the order of > 1 m when using secondary GCP distribution *B*, see section 5.3.1.4.

5.2.5. Deriving a DEM-of-Difference

Comparison of the TLS (reference) model and the SfM reconstruction were facilitated by generating DEM of Difference (DoD), based on the following DEM production process. The decimated point clouds were first aligned using cloud-to-cloud registration based the Iterative Closest Point algorithm implemented in Cloud Compare (Besl and McKay, 1992; Lague et al., 2013). This approach mitigates any major systematic differences between the two models that may have originated from differences in geo-registration. Following the approach outlined in Chapter 3, the registered point clouds were then interpolated to a TIN and linearly resampled to produce a raster DEM at 0.002 m. The DEMs for comparison against the control TLS dataset were obtained as follows:

5.2.6. Experiment A: Determining the optimum dual sensor configuration

A total of 80 images were captured comprising the following viewing geometries: nadir (n = 32), 15^{o} aft-nadir (n = 16), 30^{o} aft-nadir (n = 16) and 45^{o} aft-nadir (n = 16). These images were then used to compile a range (scenarios) of photogrammetric networks with different combinations of viewing angles as described in Table 5.2.

Table 5.3 The individual angle components of each imaging scenario 1-4 and corresponding total image numbers. All images from Angle 2 were captured from the perimeter transects to guarantee coverage of the flume.

Scenario	Angle 1 (n)	Angle 2 (off nadir) (n)	Total (n)
1	90 (32)	N/A	32
2	90(32)	15(16)	48
3	90(32)	30 (16)	48
4	90(32)	45(16)	48

Scenario 1 represents the nadir-only reconstructions of the flume, whilst scenarios 2-4 represent convergent (nadir and aft-nadir) reconstructions. The angle interval between

the oblique aft-nadir imagery is set to 15^{o} for two reasons; (1) the aft-nadir angle of 15^{o} is proposed by Wackrow and Chandler (2011) and; (2) the relative inaccuracy of the ball joint restricts the testing of angle variations below this 15^{o} threshold of change. Oblique imagery was captured along the outer transects from the same imaging location as the nadir orientated cameras to simulate the dual sensor rig proposed for the field campaign. Each convergent image set, for example scenario 3, therefore comprises 32 images with a nadir orientation and 16 oblique images acquired on the perimeter transects at 30^{0} off-nadir.

5.2.7. Experiment B: Imaging density experiments

As described later, the optimal viewing geometry was found to be based on a 30^{o} offset, and this geometry was therefore used for later experiments. This experiment focused specifically on the effect of image density which was analysed by systematically removing images at consistent camera locations for both nadir and convergent reconstructions. This approach gives rise to seven image sets as described in Table 5.4.

Table 5.4 Image number testing configurations for both nadir and convergent scenarios. Imaging locations *L* corresponds to the number of unique sensor positions upon the frame mounted roller, whereby both convergent and nadir may be captured. Nadir total plus oblique total provides the convergent total. See figure 5.4 for spatial distribution within the flume.

Imaging Locations (*L*)	**8**	**12**	**16**	**20**	**24**	**28**	**32**
Rows used	3,6	1,4,8	1,3,5,7	1,3,4,6,7	1,2,4,5,7,8,	1,2,3,4,6,7, 8	1-8
Nadir total image numbers	8	12	16	20	24	28	32
Naming convention	**8Nad**	**12Nad**	**16Nad**	**20Nad**	**24Nad**	**28Nad**	**32Nad**
Oblique image numbers	4	6	8	10	12	14	16
Convergent total image numbers	12	18	24	30	36	42	48
Naming convention	**12C**	**18C**	**24C**	**30C**	**36C**	**42C**	**48C**

The process of removing images was achieved by removing rows on imaging locations, whereby the sparsest selection comprised just two rows. Row distribution was chosen to allow sufficient overlap to successfully reconstruct within APS. The nadir image set, for example, incorporating 12 camera locations, consisted of 12 nadir images from rows

1, 4 and 8. In contrast convergent imagery from 12 imaging locations consisted of 12 nadir images and 8 oblique, from rows 1, 4 and 8 (Figure 5.4).

5.2.8. Experiment C: Imaging format

The TIFF imagery used in experiments A and B was compressed using Nikon Capture NX 2 to generate images in a lossy JPEG format. This form of compression involves the use of a discrete cosine transformation which acts like a Fourier Transform to eliminate frequencies within an image that are typically difficult for the eye to distinguish. This further reduces the information available for each pixel by rounding digital numbers to the nearest integer and ultimately reduces pixel fidelity. Such compression tends to focus on the retention of contrast over colour, enabling the identification of features (typically detected as regions of contrast). The compression reduced the mean image size from 40 MB to 3 MB, reflecting the homogenous nature of large sections of the modelled channel bed. Reconstructions were then performed in line with Experiment A - scenarios 1 and 3. The resulting models were then assessed against reconstructions from the TIFF imagery using a DoD approach

5.3. Flume experiment results

5.3.1. Experiment A: Optimum dual sensor configuration

5.3.1.1. Overview of reconstructions within APS

The initial reconstruction statistics for each scenario are presented in Table 5.5. There is a significant variation in the number of sparse points between nadir and convergent reconstructions, although this is less evident following densification. Reprojection error, which relates to poor localization accuracies between correspondences (see section 3.4), is shown to be increased under convergent imagery. There is also a non-linear increase in the processing time, as a 50% increase in image numbers for scenarios 2-4 has resulted in a 100% increase in processing times.

Table 5.5 Reconstruction statistics for each scenario available from APS where processing time relates to the sparse cloud generation.

Scenario	Cameras Aligned	GCPs	Processing Time (mins)	Sparse Points	Dense Points	Reprojection Error (Pix)
1	32	4	6:00	87,000	3.5×10^7	0.55
2	48	4	12:00	110,000	3.7×10^6	0.61
3	48	4	9:30	117,000	3.6×10^6	0.60
4	48	4	10:45	110,000	3.2×10^6	0.59

5.3.1.2. Optimization Process within APS

Remote Alpine environments present a hostile working environment that is not conducive to setting a dense distribution of ground control. The use of convergent imagery is therefore proposed to both improve reconstruction accuracy and reduce the subsequent reliance upon ground control. This reliance is most keenly felt during the optimization process, in which interior and exterior pose are re-estimated using the GCP network as exterior constraints. The ability of the optimization phase to correct the final reconstruction has been observed previously in Chapter 4. In this instance initial camera misestimation resulted in changes of up to 80 m being applied during the optimization phase. In contrast, improved pose estimates resulted in changes of just 8 m. Despite this disparity in correction required, the final estimated camera positions post optimization are of a mean displacement of just 2 m between reconstructions.

This encouraging final result, however, masks the significant corrections applied to both the pose and interior camera calibration, which propagate into the estimation of less well constrained regions of the model without ground control. A first step in understanding the value of incorporating convergent imagery therefore is the analysis of the optimization process within APS. This can be achieved by comparing DEMs derived from the sparse cloud generated before and after the optimization, using a standard DEM Differencing strategy. This analysis provides an insight into the degree of distortion and its subsequent correction. The results obtained for each experimental scenario are shown in Figure 5.5.

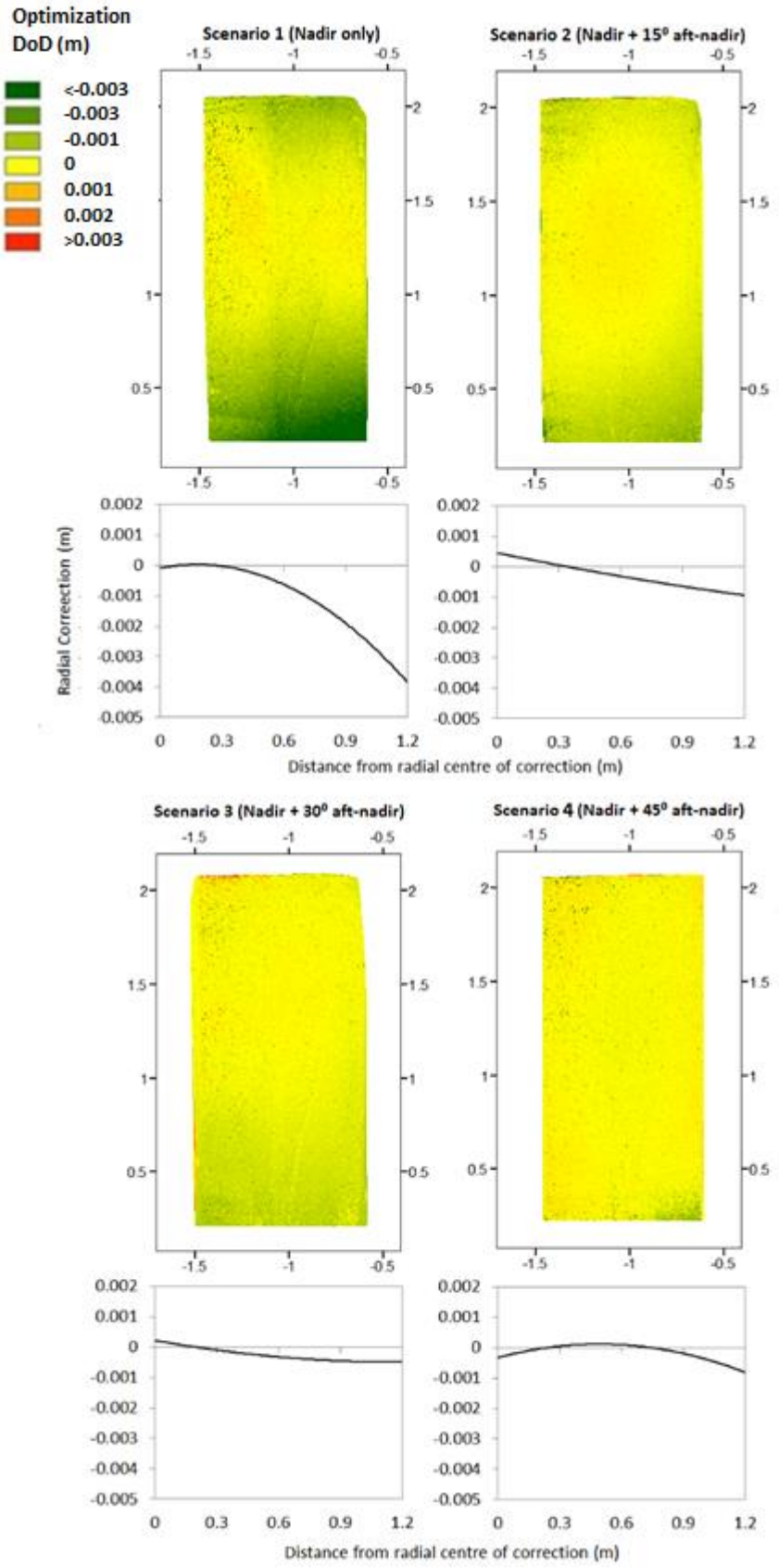

Figure 5.5 Dem-of-Differences between each angle scenario, where DEM 1 – DEM 2 equates to post-optimization minus pre-optimization. This reflects the lens correction model that is applied in APS dependent upon information derived from the GCPs. Graphs represent the gradient of correction applied when potted along a transect from the radial centre of correction.

The nadir only image network (Scenario 1) appears to show largest correction during optimization, with the DoD clearly identifying the correction of a significant dome-type distortion. This correction is most significant beyond the extent of the ground control, where corrections exceed >0.003 m. Towards the centre of the photogrammetric block, only limited correction is apparent. In contrast, optimization correction apparent in all of the convergent camera network scenarios (2, 3 and 4) appears to be much more moderate. Scenario's 3 and 4 in particular are subject to the least correction, although scenario 4 appears to display an increase in surface noise, potentially due to the very oblique camera angles leading to poor tiepoint identification and correlation.

These results suggest that scenario 3 (incorporating nadir and 30^{o} aft-nadir perspectives) provide the most robust initial bundle adjustment solved using inner constraints only (i.e., before optimization). This finding is consistent with the sensor to sensor angle proposed by Wackrow and Chandler (2008); albeit this survey is of a nadir/oblique configuration as opposed to oblique only.

The nature of correction can be examined further by plotting a transect of vertical displacement against distance from the radial centre of the photogrammetric block as also shown in Figure 5.5. These transects clearly reveal the strong correction applied to the nadir only camera network. Here, the correction is some four times the magnitude of that found for scenario 3 at a distance of 1.2 m, and over twice that of scenario 4. The transects also clearly depict the non-linearity of the correction of the nadir imagery and show the distortion accelerating in areas poorly constrained by ground control (beyond 0.6 m). By contrast, the convergent scenarios show more moderate correction and in the case of scenarios 2 and 3 a more linear pattern of correction with distance. Scaling these results to the field prototype would suggest an equivalent vertical correction of the order of 1 m displacement at a distance of 300 m from the centre of the image block, compared to just 0.25 m for scenario 3.

The impact of the optimization process can also be revealed by examining the change in the lens model parameters in Table 5.6.

Table 5.6 Radial distortion and Focal length parameters for each scenario, for both pre and post optimization within APS. Focal length is presented in pixels, while K1 is reported in units of Focal length.

Parameter	Scenario 1	Scenario 2	Scenario 3	Scenario 4
Radial distortion (k1) pre-optimization	-0.127359	-0.120927	-0.122162	-0.122229
Radial distortion (k1) post-optimization	-0.129314	-0.122903	-0.121868	-0.122215
Focal length pre-optimization (pix)	7399.69	7404.18	7394.29	7378.04
Focal length post-optimization (pix)	7538.45	7382.24	7386.31	7375.32

The effect of optimization is clearly evident on the primary radial distortion parameter, K1. The percentage change in the parameter before and after optimization is 1.54%, 1.61%, 0.24% and 0.01% respectively for scenarios 1 -4. There is a clear reduction in change required by scenarios 3 and 4, however scenarios 1 and 2 are similar. The percentage adjustment experienced by the focal length (1.88%, 0.30%, 0.11% and 0.04%. respectively for scenarios 1-4) displays a clear disparity between all of the convergent scenarios this time. In this instance the pattern observed in Figure 5.5 appears more closely related to the focal length than radial distortion. Between the surveys, the final parameters can be seen to vary between scenario 1 (K1 = -0.129314, Focal Length =7538.45) and 3 (K1 = -0.121868, Focal Length =7386.31) by 5.75% and 2.02% respectively. This indicates that despite the use of high quality ground control to constrain the final stage of the bundle adjustment within APS, there are inherent structural differences between scenario 1 and 3 as a result of differing calibration estimates.

The flume further provides the opportunity to assess the impact that changing calibration may have upon the 32 cameras that exist both within the nadir only and convergent scenarios. These 32 nadir orientated images are identical across all scenarios yet as part of the global bundle adjustment result in differing calibration estimates between scenarios. This change is expressed primarily in the exterior orientation of the camera. This is due to the significance placed upon the ground control within the bundle

adjustment which results in differences in camera calibration manifesting in the exterior orientation. Once more using scenario 1 and 3 as an example, the final camera locations are observed to be systematically offset at a vertical magnitude of 0.03 m which corresponds to ~2.5% of the camera height.

This result corroborates the findings of Wackrow and Chandler (2008) who proposed that convergent imagery provided improved depth estimations. In this situation, the bias of the bundle adjustment towards the high quality ground control coupled with poor internal calibration estimation necessitates compensation to the simultaneous solution through adjustments to the exterior orientation. This interaction between calibration estimates and pose further the findings from Chapter 4, whereby incorrect camera positional estimation lead to deterioration in model accuracy. This was seen during the optimization phase as internal parameters required significant alterations to overcome pose estimation of >80 m

5.3.1.3. SfM-TLS Comparison

The analysis described above reveals key differences between the scenarios, but does not provide an indication of the accuracy of the reconstructions. This can however be achieved readily through comparison with the TLS model, here taken as an independent reference surface. Such an analysis is crucial, for it was clear in Chapter 4 that where ground control is well distributed, post-hoc optimization can generate effective results. It is important therefore to ensure that convergent imaging does not lead to a deterioration of the model, both within areas of good GCP control and beyond.

Figure 5.6 therefore shows DoDs comparing each of the four scenarios described above to the TLS reference model. Here the DoD is colour rendered, to show areas where the TLS is vertically above the SfM surface in shades of red (+ve values) and in green where the TLS is below the SfM model. The frequency distribution of elevation changes is also shown for each scenario.

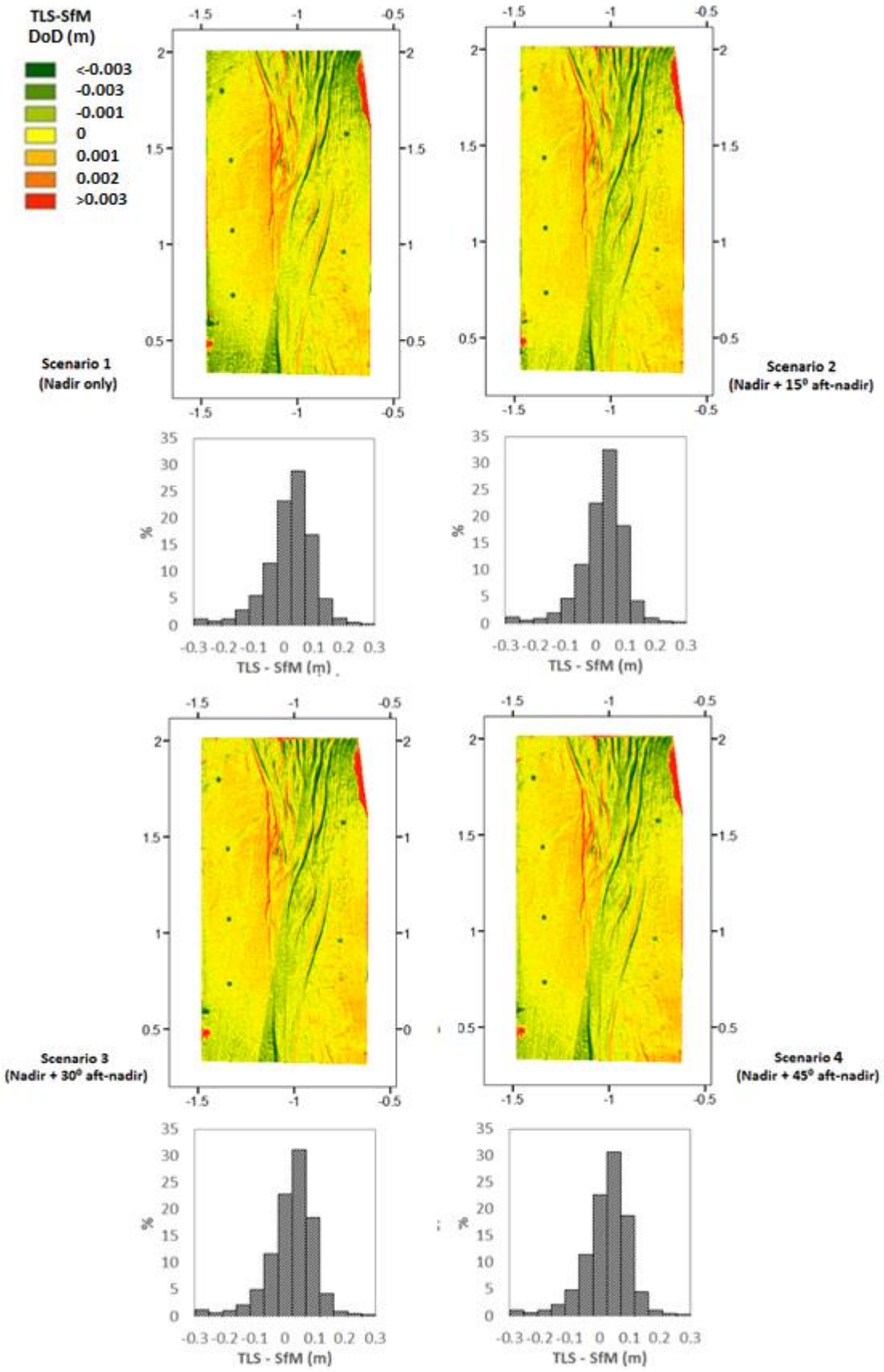

Figure 5.6 Dem-of-Differences between the TLS reference dataset and Scenario 1-4 reconstructions, where DEM 1 – DEM 2 equates to TLS – SfM. A histogram of elevation residuals is plotted.

The effect of the dome shaped distortion is evident in the nadir image scenario 1, with the TLS model above the SfM reconstruction in the centre of the block and lower around the margins. For illustration, the bottom left of the scenario 1 DoD indicates up to 0.002 m lowering, a pattern observed in the top left and right corners too. Scaling these values to the field result in a 0.5 m vertical error at a distance of 250 m from the closest GCP. In contrast the degree of lowering at the margins for all convergent scenarios (1-4) is reduced. Scenario 3 experiences just a 0.001 m lowering at the bottom left margins

However, the optimization procedure has produced an overall surface geometry that is broadly comparable to the convergent models shown by the DoDs for the remaining scenarios. A summary of errors are presented in Table 5.7.

Table 5.7 Key error metrics relating to each scenario, 1) Nadir, 2) Convergent (nadir + 15º aft-nadir), 3) Convergent (nadir + 30º aft-nadir) and 4) Convergent (nadir + 45º aft-nadir).

Scenario	Area <0.002 m	Area <0.0005 m	LOOCV error (m)	ME (m)	MAE	SDE (m)	MaxE (m)
1	93.75%	51.08%	0.0163	0.0002	0.0008	0.0020	0.0309
2	94.35%	53.84%	0.0152	0.0002	0.0008	0.0019	0.0305
3	94.41%	52.81%	0.0136	0.0002	0.0008	0.0019	0.0304
4	94.42%	52.27%	0.0134	0.0002	0.0008	0.0019	0.0304

The percentage of the DoD showing differences of less than 0.002 m and 0.0005 m representatively are broadly similar across all scenarios. A more discriminating measure of difference appears to be the Leave One Out Cross Validation (LOOCV) analysis of based on sequentially removing one GCP from the reference set and evaluating and accumulating the error (see Chapter 3). The LOOCV error is smallest in both scenarios 3 and 4, with values of 0.0136 and 0.0134 m respectively, displaying a high level of consistency between the convergent camera networks. The largest error is observed for the nadir orientated model (0.0163 m), some 21% higher than scenario 3.

Despite the diffuse lighting conditions, reconstruction of the channel margins appears to cause some problems for all scenarios. This possibly reflects the impact of poor tiepoint identification and matching, but could also relate to interpolation artefacts in

areas of comparatively sparse data. Improvements in surface geometry resulting from convergent imagery are constrained to the peripheral regions of the flume beyond the extent of the ground control. Areas within the perimeter of the control network appear to generate comparable results, regardless of the camera network. The magnitude of variation is relatively small across all error metrics, potentially as the scale of deformation is not significantly larger than the basic error in measurements obtained from the SfM, TLS and Total station.

5.3.1.4. Impact of varying the control network

With the use of the primary GCP distribution (TL2, TL3, TL4 and TR2 in Figure 5.4), the relative extent of the flume area outside of ground control is fairly minimal, dampening the scale of improvements offered by convergent imagery. Results from Chapter 4 suggest that inaccuracies become significant at a distance beyond the GCP perimeter approximately equal to the GCP radius diameter. To address this, scenarios were reconstructed using a reduced set of ground control aimed at exaggerating the degree of error recorded at a number of independent control points (ICPs) see Figure 5.7.

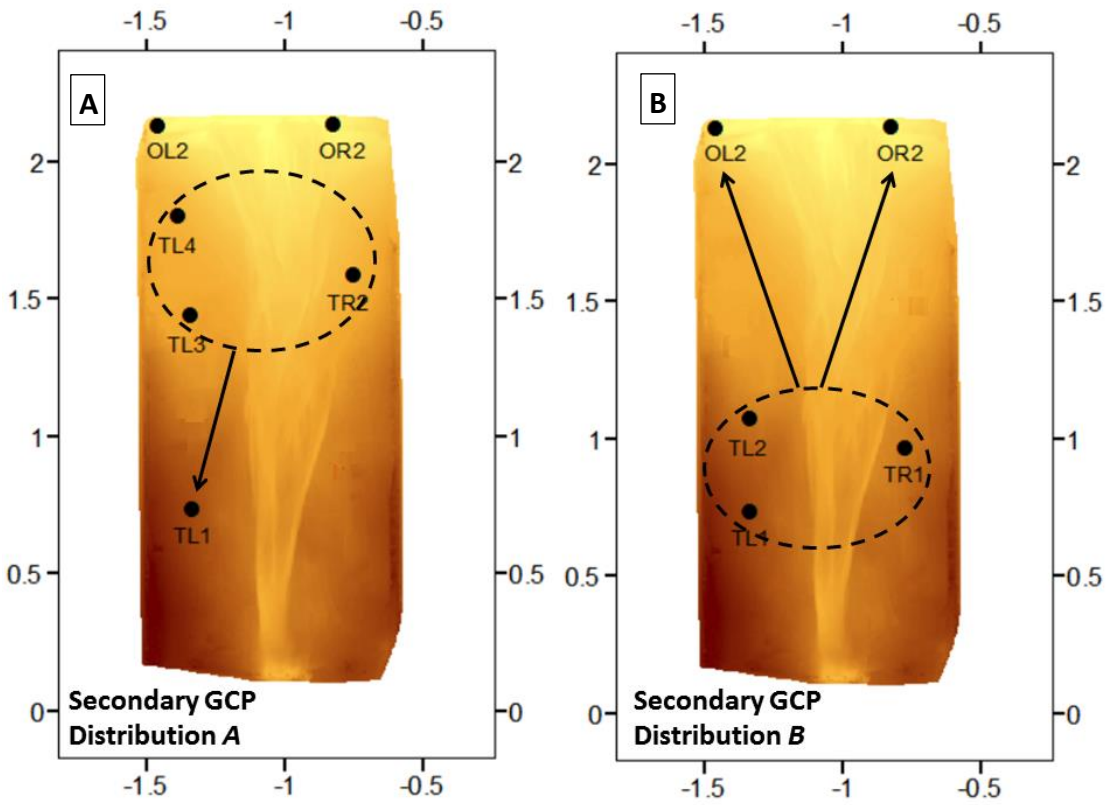

Figure 5.7 The secondary GCP scenarios derived to exaggerate ICP errors. In this instance the marker TL1 provides an ICP value for distribution *A*, whilst OL2 and OR2 provide an ICP value for distribution *B*.

By increasing the distance between the perimeter of ground control and the ICP chosen, it is possible to exaggerate the error value that is recorded. This potentially provides a greater insight into the degree of systematic error that is present in each scenario that may have been masked by the strong ground control network offered by the primary GCP distribution.

Table 5.8 Errors recorded for ICPs from the reconstructions undertaken with secondary GCP distributions aimed at exacerbating the error values recorded. For marker reference see Figure 5.4.

Marker	Secondary GCP distribution	Distance to Nearest GCP (m)	Scenario 1 error (m)	Scenario 2 error (m)	Scenario 3 error (m)	Scenario 4 error (m)
TL1	A	0.7	0.010	0.009	0.008	0.004
OL2	B	1.0	0.010	0.009	0.006	0.006
OR2	B	1.2	0.011	0.008	0.006	0.008
Mean		**1.0**	**0.010**	**0.009**	**0.007**	**0.006**

The analysis of independent check points under these exaggerated conditions in Table 5.8 highlights the presence of increasing systematic error in the nadir orientated models, with an average increase in error of ~50% in comparison to convergent scenarios 3 and 4. Scaling these deviations to the field would result in deviations of 20 m at a distance of 300 m from the perimeter of ground control. This exacerbated error, which fails to correlate with field measurements, is resultant from the minimal number of GCPs (n = 3) present in the secondary GCP distributions highlighting the importance of recording abundant control. However, the process is able to determine a marked improvement and reduction in error propagated through insufficient ground control distribution through the adoption of convergent scenarios.

5.3.1.5. Summary: The impact of varying imaging angles

The results presented above indicate that nadir viewing geometries can yield faithful models when supported by dense ground control and post-hoc camera optimization. However, these models appear to suffer significant distortion without ground control, or (as would be more typically observed) in areas beyond the extent of any control network. By contrast, convergent image networks appear to generate more robust

solutions to the initial bundle adjustment and are therefore less reliant on post-hoc camera calibration. This result is of particular significance for large scale surveys to be undertaken here on the Dart River given the inevitable sparse coverage of control that can be practically achieved.

The camera network described in scenario 3 is identified as the most suitable configuration for a dual camera rig, comprising two sensors offset at 30^{o} (nadir and 30^{0} aft-nadir). The imaging angle configuration is prone to less noise in the optimization phase and generates reliable results as shown by comparison against the TLS reference model. Increasingly the obliquity of images as described in Scenario 4 (angle of 45^{o} aft-nadir) appears to generate large pixel distortions that either fail to match or result in erroneous tiepoints. While future studies will therefore focus on camera separation angles of only 30^{o}, it is important to note that the aerial survey is likely to incorporate such low angle photography during the turning circles of the helicopter at the end of flight lines. From this point on in this chapter the phrase nadir imagery relates to the survey scenario 1, whilst convergent imagery refers to survey scenario 3.

5.3.2. Experiment B: Optimum capture frequency

5.3.2.1. Computational impact

Adopting a secondary camera in future surveys, as suggested by the previous analysis will lead to a doubling of imagery captured. This presents severe computational implications for catchment scale studies involving large image sets. Determining the optimal minimum image density (i.e., image overlap) within APS could potentially lead to an increase in the efficiency of data collection through reduced field survey time and the subsequent processing of images. This is particularly relevant for studies in which large image sets may be constrained by the amount of computational power available and may restrict reconstructions to the use of lower alignment settings within APS.

Here, varying image density has been achieved not by taking further photography, but by systematically thinning the image sets acquired for Experiment A. In this, rows of images are removed to alter baseline distances and in so doing, alter the image overlap see Table 5.4; see Figure 5.4 for reference. Due to the nature of how this imagery has been thinned, a constant sidelap has been maintained throughout which relates to the

distance between imaging transects. From this point forward the image density shall be defined in terms of the overlap percentage as is standard within the literature. This value has been quantified by averaging the endlap achieved at each baseline distance tested with the constant sidelap value of 71%.

This thinning process has been applied to both nadir and convergent imagery, specifically scenarios 1 and 3, to assess their robustness to varying overlap. Reconstruction statistics from processing within APS are presented for each image density in Table 5.9. Imaging location, *L*, relates to the discrete number of location on the frame mounted roller above the flume that imagery was captured from, where the suffix Nad represents the nadir imagery of scenario 1 and the suffix C represents the convergent network of scenario 3.

Table 5.9 APS reconstruction statistics for all image density scenarios. 8Nad represents 8 nadir orientated images and 8C (12) represents 12 convergent images from 8 image locations referred to as *L* in figure 5.8. * Values for overlap are not presented for the convergent imaging network. The footprint of nadir orientated imagery falls broadly within the flume surface and as such is easy to compute, whereas the aft-nadir imagery consists of a reduced and varied percentage of footprint within the flume, complicating the calculation.

Imaging Locations (*L*)	**Cameras Aligned**	**Overlap**	**Time (minutes)**	**Sparse Points**	**Dense Points**	**Reprojection error (pix)**
8Nad	8	36%	0:50	36,000	2.5×10^6	0.49
8C (n = 12)	12	*	1:20	48,000	2.9×10^6	0.53
12Nad	12	52%	1:30	51,000	3.2×10^6	0.50
12C (n = 18)	18	*	2:20	66,000	3.4×10^6	0.57
16Nad	16	69%	2:05	65,000	3.3×10^6	0.52
16C (n = 24)	24	*	3:45	85,000	3.6×10^6	0.59
20Nad	20	73%	2:50	71,000	3.3×10^6	0.53
20C (n = 30)	30	*	4:30	94,000	3.6×10^6	0.59
24Nad	24	75%	3:50	80,000	3.5×10^6	0.53
24C (n = 36)	36	*	6:00	106,000	3.8×10^6	0.60
28 Nad	28	77%	4:30	84,000	3.5×10^6	0.55
28C (n = 42)	42	*	7:50	113,000	3.8×10^6	0.59
32Nad	32	80%	6:00	87,000	3.5×10^6	0.55
32C (n = 48)	48	*	9:30	117,000	3.6×10^6	0.60

This analysis shows that reprojection error increases with both convergent and nadir imagery and as the size of the image set grows. The most relevant statistic revealed here however is the increase in processing time which grows systematically with the size of the image set, as shown in Figure 5.8.

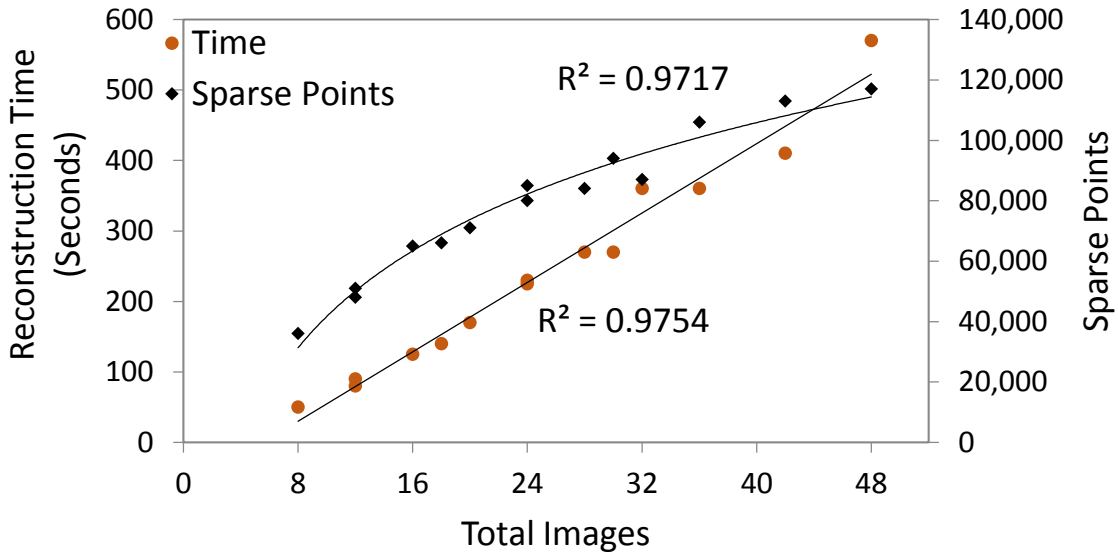

Figure 5.8 Regression of reconstruction time and sparse point cloud numbers within APS against image numbers. The plot for total image numbers contains both nadir and convergent network scenarios.

Under the disabled alignment function within APS a nonlinear increase in processing time is to be expected. However, the processing time in Figure 5.8 obeys a linear pattern of increase with an r^2 value of 0.97. This may be due to the limited sample size (n = 7) from which to derive information or indicate that the alignment time within APS is also strongly related to the number of points within the sparse cloud. The relationship between processing time and image numbers is non-linear and suggests that past a threshold of image density, further imagery serves primarily to increase processing times with the introduction of minimal additional tiepoints. Regardless, the impact of increasing image numbers upon computational time is clear and even under a linear increase would indicate a doubling of processing time with a dual camera rig.

5.3.2.2. SfM-TLS Comparisons

It is important to consider the impact image density has upon the accuracy of reconstructions, through benchmarking against the TLS data (Figures 5.9 and 5.10). The impact of reducing image density is likely to be most acute at the margins of the model given the partial coverage in these areas. This is evident across all reconstructions where image density at the centre of the block often bears no resemblance to that at the margin.

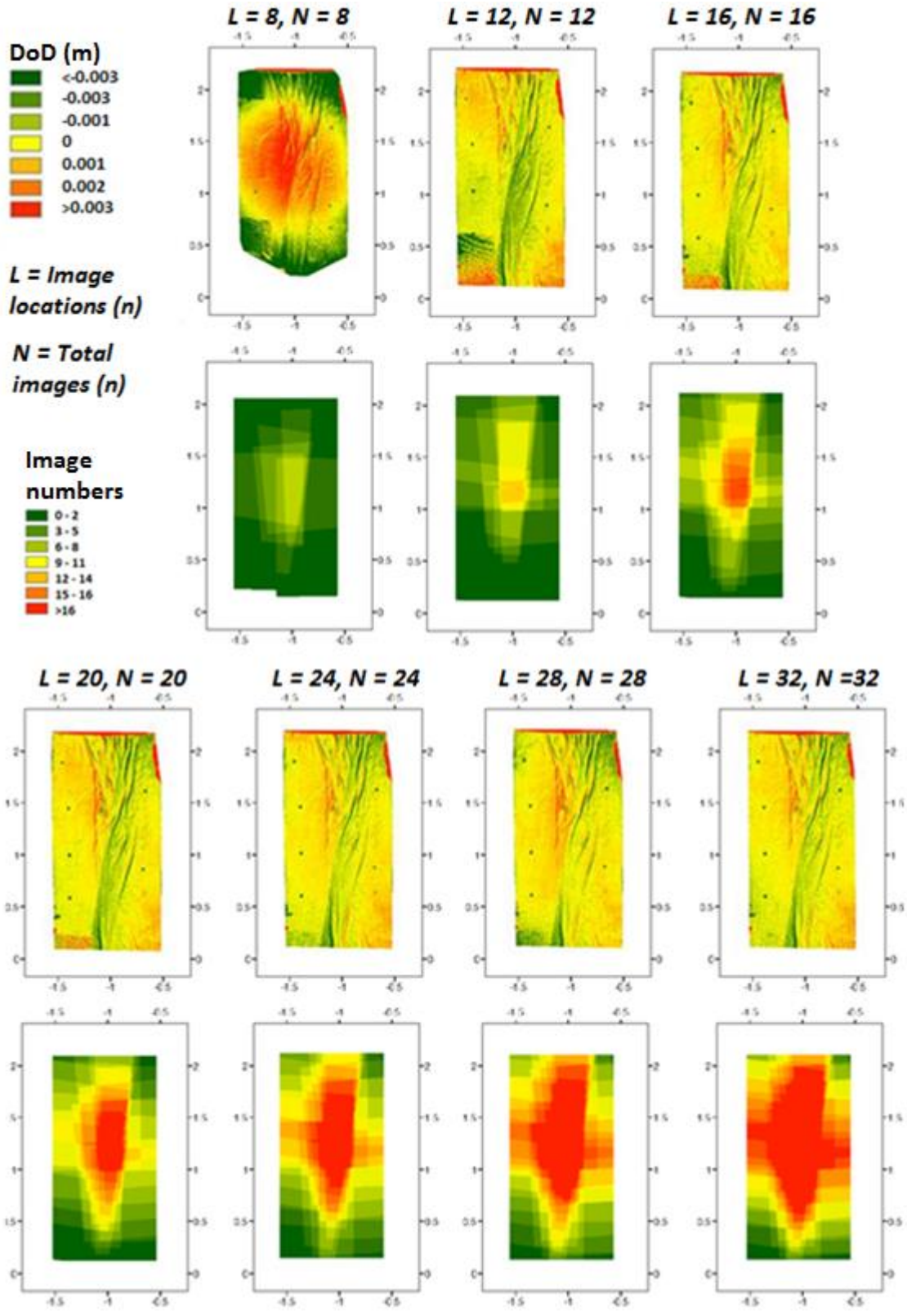

Figure 5.9 TLS – SfM DoD and corresponding map for image density for nadir only imagery of the flume. Imaging locations *L* are consistent for both nadir and convergent reconstructions.

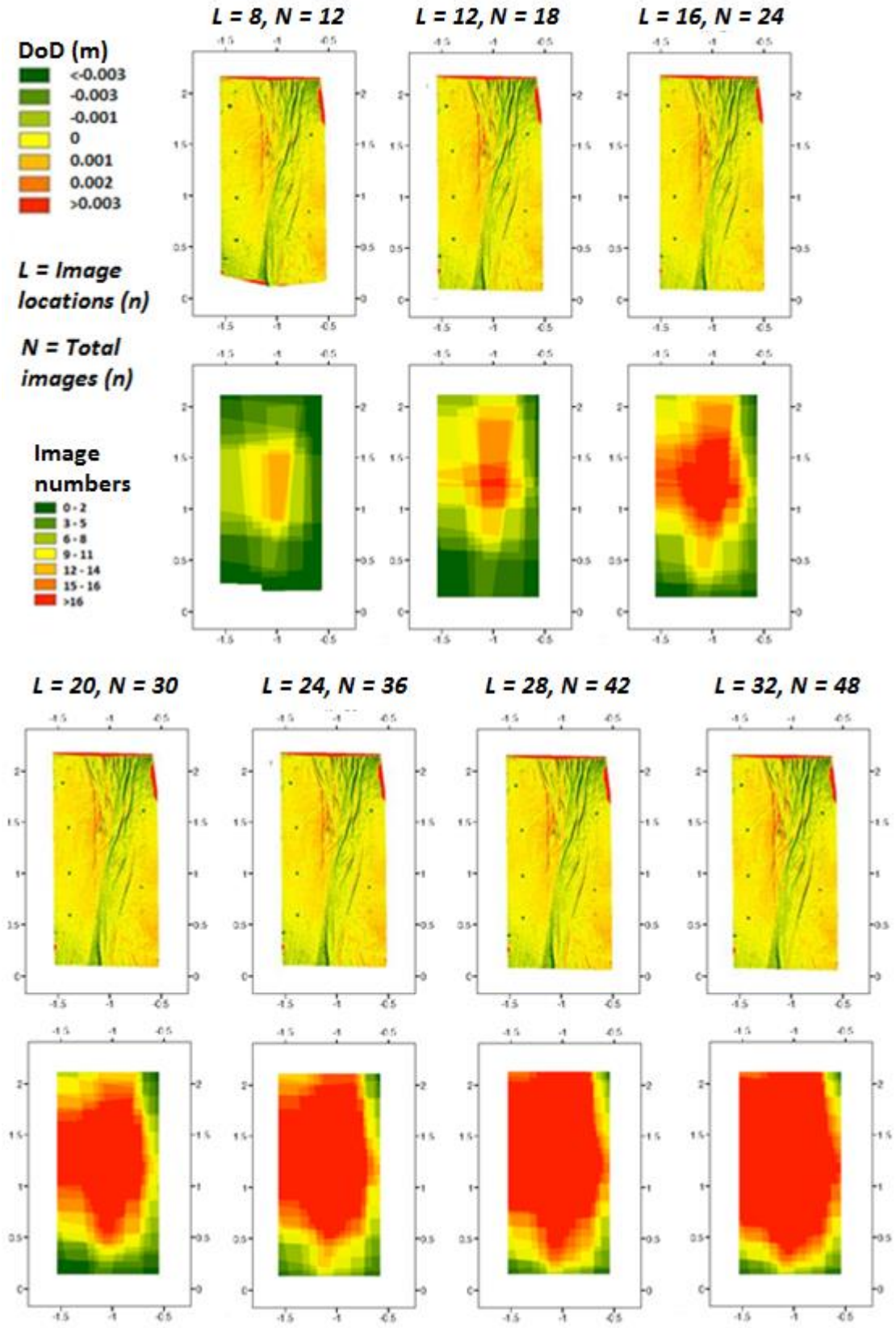

Figure 5.10 TLS – SfM DoD and corresponding map for image numbering for convergent imagery of the flume. Imaging locations *L* are consistent for both nadir and convergent reconstructions.

Table 5.10 Image location and total numbers with corresponding error metrics. For example, where *L* = 8, within the total image column 8 refers to the nadir scenario and 12 (C) the convergent equivalent.

Image Locations (*L*)	Total Images (N)	Mean Error	MAE	SDE (m)	LOOCV (m)	Max Error	Area <0.002 m	Area <0.0005 m
8	8	0.0001	0.0020	0.0028	0.0083	0.0306	62.80%	16.75%
	12 (C)	0.0002	0.0008	0.0020	0.0093	0.0304	94.57%	53.48%
12	12	0.0001	0.0009	0.0020	0.0152	0.0316	91.11%	43.38%
	18 (C)	0.0002	0.0008	0.0019	0.0106	0.0305	94.72%	53.86%
16	16	0.0002	0.0009	0.0020	0.0137	0.0319	93.49%	52.67%
	24 (C)	0.0002	0.0008	0.0019	0.0133	0.0303	94.90%	55.13%
20	20	0.0002	0.0008	0.0020	0.0108	0.0316	93.76%	50.74%
	30 (C)	0.0002	0.0008	0.0019	0.0165	0.0303	94.99%	54.11%
24	24	0.0002	0.0008	0.0019	0.0049	0.0316	94.67%	52.91%
	36 (C)	0.0002	0.0008	0.0019	0.0181	0.0302	94.94%	53.95%
28	28	0.0002	0.0008	0.0019	0.0047	0.0304	93.88%	51.95%
	42 (C)	0.0002	0.0008	0.0019	0.0073	0.0303	95.05%	53.96%
32	32	0.0002	0.008	0.0020	0.0151	0.0309	93.75%	51.08%
	48 (C)	0.0002	0.0008	0.0019	0.0095	0.0304	94.23%	52.81%

The nadir n = 8 model shows significant doming deformation following a concentric ring pattern, furthermore the percentage of area <0.002 m is seen to vary from 62.8% in contrast to ~90% observed in all other reconstructions (see Table 5.10). This is likely to be an impact of the reduced image numbers limiting GCP identification, with just three projections observed for each GCP. The number of projections from which GCPs are observed is specifically highlighted within the APS manual as being related to model accuracy (Agisoft, 2017). This doming effect is essentially removed from n = 12 onwards, however, there is evidence of image artefacts in the bottom corners where the reduced overlap has led to localized areas of poor reconstruction quality. At n = 24 there are no longer any obvious local artefacts, but also no further improvement in reconstructed

geometry. This is clearly evidenced by a LOOCV (see Table 5.10) analysis, which shows a systematic reduction from 0.0152 m (n = 12) to 0.0137 m (n = 16) and 0.0108 m (n = 20) and then remaining similar at 0.0049 m (n = 24) and 0.0047 m (n = 28). Furthermore, area of the model deviating from the TLS by more than <0.0005 m is increased from 16.75% (n = 8) to 43.38% (n = 12) before becoming stable at n = 16 and beyond with 53.86% <0.005 m.

The DoDs for the convergent models reveal a significant improvement over the nadir models at lower imaging numbers, whilst high imaging numbers reflect the continuation of improvements observed in 5.3.1. There are no artefacts visible within the DoD analysis, however, the bottom left corner of n = 12 fails to reconstruct. The error metrics table would suggest that once n = 18 is achieved results improve no further, with 94.72% < 0.002 m and 53.86 m < 0.0005 m and a LOOCV of 0.0106 m. This is compared to the reconstruction using full convergent imagery (n = 48), of 94.23% < 0.002 m, 52.81% < 0.0005 m and a LOOCV of 0.0095. The image density achieved by n = 18 would correspond to an overlap of just 52% as recorded by standard singular image capture strategies. This however is a dual rig and as such would equate to an overlap of 72% when averaged over a sufficiently large area.

When analysing the metrics in table 5.10 there are no systematic trends between nadir and convergent imagery aside from the percentage of area within 0.002 m and 0.0005 m, of which convergent imagery is consistently improved. The most significant difference between convergent and nadir models is instead observed spatially in the bottom left region of the flume, where systematic distortion is present. During the image density experiments this area was subject to image artefacts for nadir n = 8, 12, 16 and 20, despite having the same coverage in these regions as the convergent reconstructions. The overlap threshold for nadir imagery would therefore appear to consist of a significantly smaller baseline distance. As a result it may be seen that convergent networks under a dual rig scenario are robust to greater baseline distances, which may be facilitated within the field through greater flight speeds.

When reconsidering the computational cost of image numbers within APS it may be seen that n = 18 under a convergent dual camera rig results in an improved reconstruction against all nadir scenarios, including n = 32 and an equivalent to convergent n = 48 (table

5.10). It would appear that the introduction of additional viewpoints not only reduces the number of image locations required, but to such an extent to reduce the total number of images required full stop. In this instance, a reconstruction with 18 images produces improved results to one with 32, requiring just 38% of the processing time and equivalent results to n = 48 at just 25% of the time.

5.3.2.3. Summary: Impact of varying image density

Increasing image numbers provides improvements in model structure up to differing thresholds for both nadir and convergent imagery. Beyond these thresholds, further increases in image density (overlap) fail to generate major enhancements in model accuracy but nonetheless add significantly to computational time. This reflects the lack of new information provided by these additional viewpoints.

The choice to represent image density as an overlap percentage is motivated by the prevalent method of reporting within the literature. For the purposes of this study it is important to consider that with a constant sidelap, the change in overlap is primarily factor of endlap, which in turn is an indication of the baseline distance. This baseline distance may be altered during imagery capture within the field by either altering the capture frequency or helicopter flight speed. Results from Chapter 4, especially when assessing the impact of varying lighting conditions suggests that similar, if not greater flight speeds are required. Using the finding from this section we are able to determine the optimum frequency of capture for a proposed flight speed of 90 knots for the 2015 survey. For example, a capture frequency of 0.71 Hz would be required to achieve the threshold of 75% overlap required for nadir only orientated imagery.

Finally, it is worth considering how the survey design may be altered to ensure sufficient overlap is maintained across the area of interest. Figures 5.9 and 5.10 indicate that the perimeter of the area of interest will always be subject to sparse image densities as endlap is maintained but sidelap is diminished. In this instance, where imagery is subject to a constant sidelap of 71%, the outer 29% of imagery will only be covered by endlap and the outer 58% will have in the order of half the coverage of imagery central to the block. To mitigate this loss of coverage surveys would be encouraged to extend the flight perimeter beyond the survey area to provide sufficient maintenance of sidelap and consequentially imaging density for the area of interest.

5.3.3. Experiment C: Impact of imaging format

5.3.3.1. Computational impact

The experiments described above suggest that convergent camera networks provide a more robust geometry for solving scene structure and can do so with lower numbers of images than a comparable nadir network. However, the deployment of a dual camera rig, capturing nadir and aft-nadir orientated photography simultaneously will double the size of the total image set. While there may be some redundancy in the resulting imagery, over capture is likely to prove useful in the first instance to mitigate unforeseen changes in the flight plan which could lead to a degradation of the acquisition strategy as happened in the Gorge section in 2014. A complementary strategy to managing this additional computational overhead is to consider whether degrading image quality rather than image numbers can achieve improved results. To assess this, here a series of experiments are described in which reconstructions based on lossless TIFF formatted imagery are compared to those obtained with lossy (irreversibly compressed) JPEG formatted images.

JPEG compression focuses on maintaining image contrast at the expense of colour, which should, at least in theory, have a limited effect on the feature detection algorithm. Such imagery may however provide significant savings in computational power, with JPEG images typically an order of magnitude smaller than uncompressed TIFF formats. To assess whether image compression has a marked impact on the estimated scene structure and dense models, Scenarios 1-4 from Experiment A are re-run here following the same approach, but using JPEG compressed images from the outset. A summary of the resulting reconstructions is show in Table 5.11.

Table 5.11 Reconstruction Statistics for JPEG imagery used within APS as compared to TIFF based reconstructions.

Scenario	Cameras Aligned	GCPs	Processing Time (minutes)	Sparse Points	Dense Points	Reprojection Error (Pix)
S1 JPEG	32	4	5:10	87,000	3.5×10^6	0.61
S1 TIFF	32	4	6:00	87,000	3.5×10^7	0.55
S2 JPEG	48	4	9:45	110,000	3.8×10^6	0.66
S2 TIFF	48	4	12:00	110,000	3.7×10^6	0.61
S3 JPEG	48	4	8:15	115,000	3.7×10^6	0.65
S3 TIFF	48	4	9:30	117,000	3.6×10^6	0.60
S4 JPEG	48	4	7:50	110,000	3.3×10^6	0.65
S4 TIFF	48	4	10:45	110,000	3.2×10^6	0.59

Once more there is large initial variation in the sparse cloud which is diminished post densification. Average reprojection errors are increased, however, the number of points within the clouds remains constant. The use of compressed JPEG imagery resulted in a 17% reduction in reconstruction time within APS for nadir imagery and 14% for scenario 3.

5.3.3.2. SfM-TLS comparison

A comparison of the JPEG derived models against the TLS reference data was achieved using the DoD approach described above. The resulting statistics reflecting the differences in the models is given below in Table 5.12. The JPEG reconstructions are comparable to those conducted with lossless TIFF imagery for Scenario 1 and 2, despite the contrast over colour approach of the JPEG compression which would be expected to lose detail in homogenous areas such as the sand-bed channel. There is however a significant increase in the LOOCV error for scenario 4. This is likely due to the loss of sharpness towards the edge of the field of view which is compounded by the image compression. This limits the use of severely oblique imagery obtained during helicopter turns.

Table 5.12 Error metrics presented for flume reconstructions in comparison to TLS when using compressed JPEG imagery (File size 3 MB vs. 30 MB). Error stats relate to DEM to DEM cell by cell calculation.

Scenario	Area <0.002 m	Area<0.0005 m	LOOCV error (m)	ME (m)	MAE (m)	SDE (m)	MaxE (m)
1 JPEG	93.88%	51.94%	0.0151	0.0002	0.0008	0.0019	0.0313
1 TIFF	93.75%	51.08%	0.0163	0.0002	0.0008	0.0020	0.0309
2 JPEG	94.35%	53.84%	0.0167	0.0002	0.0008	0.0019	0.0305
2 TIFF	94.35%	53.84%	0.0152	0.0002	0.0008	0.0019	0.0305
3 JPEG	94.23%	52.81%	0.0095	0.0002	0.0008	0.0019	0.0304
3 TIFF	94.41%	52.81%	0.0136	0.0002	0.0008	0.0019	0.0304
4 JPEG	93.83%	52.52%	0.0147	0.0002	0.0008	0.0019	0.0304
4TIFF	94.42%	52.27%	0.0134	0.0002	0.0008	0.0019	0.0304

Overall there is minimal variation in error metrics across the scenarios. A more subtle comparison is afforded by directly comparing the TIFF and JPEG SfM models directly. This can be achieved using the DoD approach again and Figure 5.10 shows DoDs for each modelled scenario but calculated based on the TIFF SfM – JPEG SfM models (Figure 5.11).

This shows that the JPEG reconstructed surfaces each scenario fall within 0.0002 m of the original uncompressed TIFF counterpart. Again, there is a clear difference between quality of the convergent and nadir camera networks. A doming pattern is observed within the nadir imagery focused at centre of the photogrammetric block. However, the reduction in keypoint fidelity may also result in an increased tiepoint error, resulting in a poorer calibration. This is best reflected by a 0.82% change in k1 parameters between the JPEG and TIFF scans. The reprojection error also increased from 0.55 pix to 0.61 pix, however, the number of points within the cloud remained constant. As the angle of the secondary camera increases, the impact of this distortion can be seen to decrease resulting in a speckled pattern, indicative of local distortions associated with a loss of feature detail. The majority of results for both 30^0 and 45^0 are confined within 0.00005 m, which would scale to a field site equivalent of just 0.013 m.

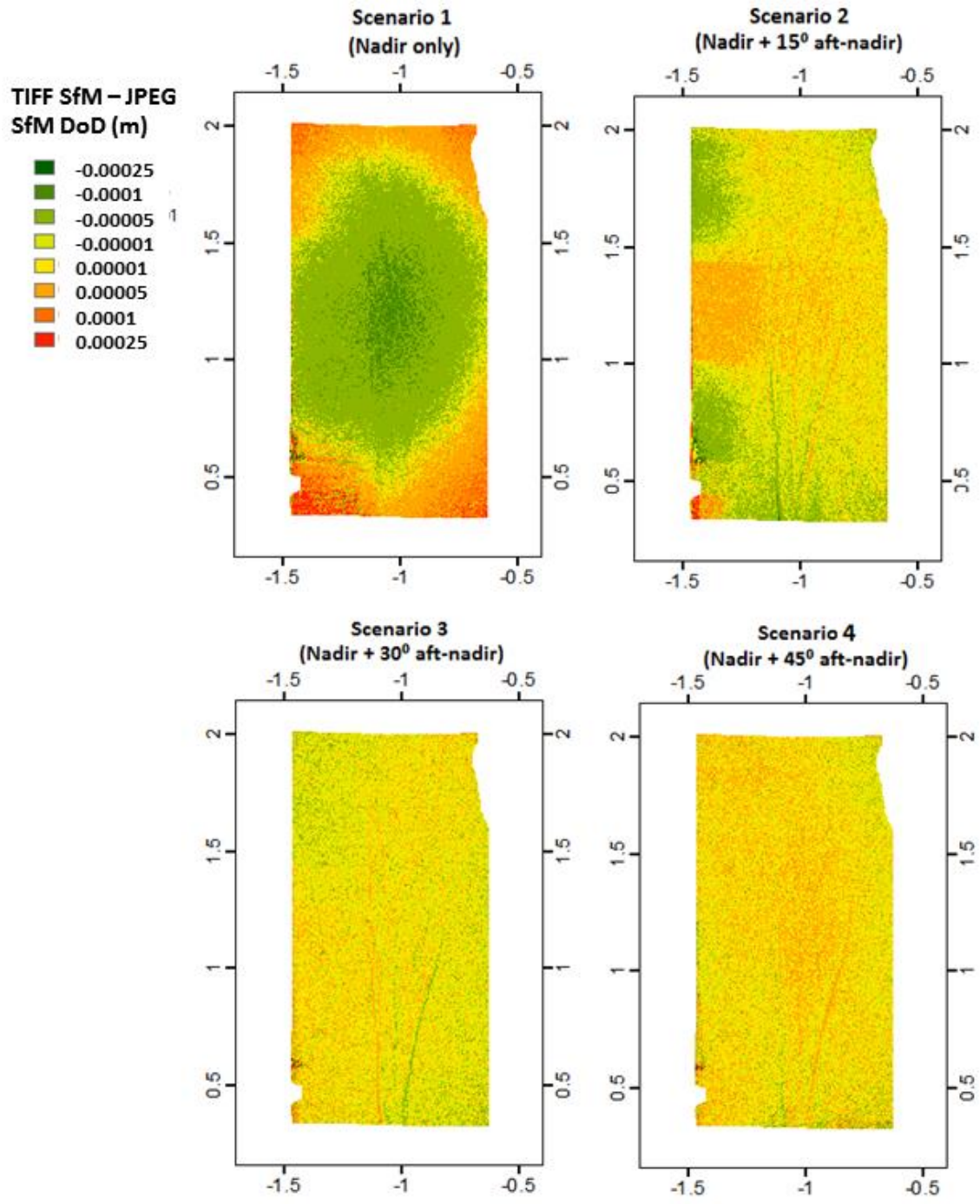

Figure 5.11 SfM-SfM DoD of TIFF – JPEG reconstructions of the flume using the sparse point cloud. Scenario 1 exhibits a non-linear disparity, however must be viewed within the context that variation is in the order of just +/- 0.0001 m.

Results would suggest that a future approach utilizing JPEG could provide a suitable strategy so long as a robust camera network is used. This provides the opportunity to use a high compression ratio (factor of 10) dramatically reducing image size. This saving

may, in part reflect the additional information gleamed from the secondary camera angle which improves feature recognition/description even with lower quality imagery.

5.4. Up-scaling to the field

The laboratory experiments described above highlight a number of strategic directions that offer potential advantages and could be adapted for later surveys on the Dart River. These include the adoption of a dual camera rig to obtain convergent photography, minimum overlap requirements for the flight planning and the use of compressed imagery which will help to reduce the computational burden. These findings have, however, all be conducted on a loosely scaled laboratory model and it is important to evaluate whether this strategy can be applied successful at the scale of the natural prototype. To evaluate this, a well-controlled field experiment is described below which uses convergent photography from a dual camera system to develop a model of the small township of Glenorchy at the top of Lake Wakatipu on the true left of the Dart-Rees delta.

5.4.1. Control point analysis in Glenorchy

This study focuses on a 1 x 1 km area covering the township of Glenorchy as shown in Figure 5.12 below. This site provides an ideal test case, given the abundance of low relief markings (i.e., road arrows, lines) that can be surveyed as ground control points and the availability of a 2011 LiDAR survey provided to the study by the Otago Regional Council. Imagery for the analysis was obtained as part of the 2015 survey using a dual camera system. This comprised two Nikon D750s (28 mm, f6.3, 1/1000s) orientated at a 30^{o} separation in a nadir-aft-nadir configuration (see Chapter 3 for details). In comparison to the Nikon D3500 used in the 2014 Dart survey and the flume experiments, The D750 has an increased field of view with an improved CCD resolution. The combination enables an increase in flight speeds and flight heights whilst maintaining the ground sampling distance and image overlap. Estimates of camera positions were also provided using an on-board code-only GNSS receiver which logs lat, long and ellipsoidal heights on the EXIF image file.

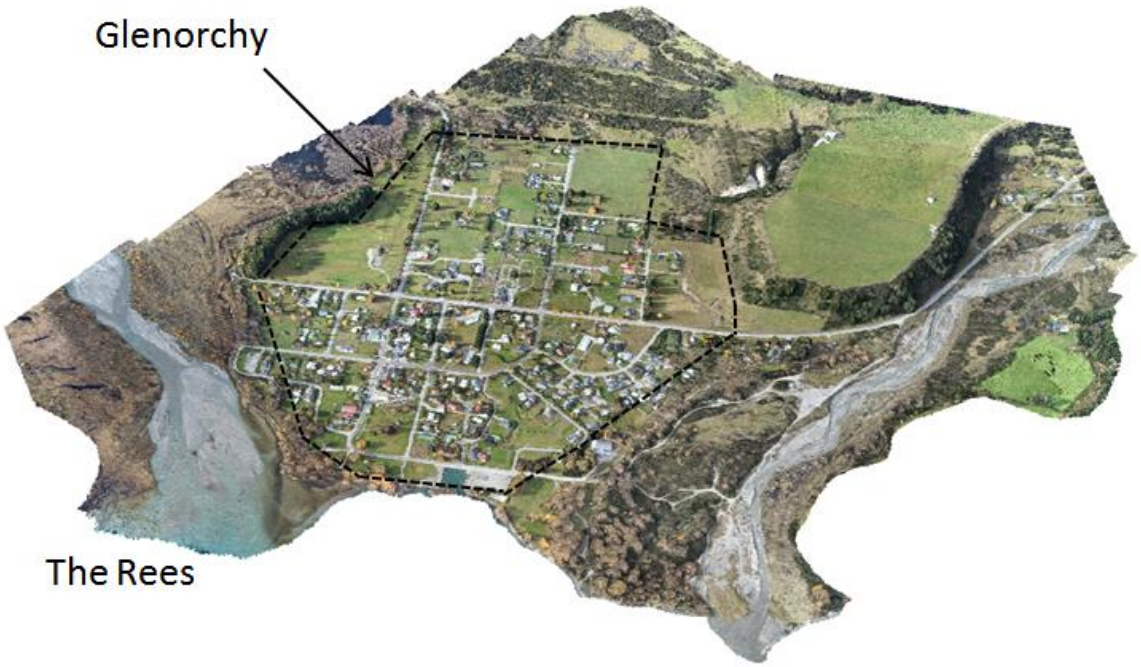

Figure 5.12 Dense reconstruction from the 2015 Dart survey of the town of Glenorchy, captured using convergent imagery. The outline corresponds to the analysis area presented in subsequent figures.

A further 27 ground control points were surveyed in the field, using clearly delimited road markings. All survey observations were acquired with RTK GNSS on a detail pole and have expected accuracies of the order of 1-5 cm. Observations were recorded in the standard NZTM coordinate system and NZGD09 to enable comparison with the 2011 LiDAR survey. The control points were subsequently subset into two groups, n_1 = all 27 points, and n_2 = 14 focused at the centre of the block (see Figure 5.12 below). The reduced set of GCPs provides the opportunity to examine how the optimization (posteriori camera calibration) process operates differentially on initial bundle adjustments based on different camera configurations (image network). The full GCP set (n =27) provides the opportunity to assess the reliance of the optimization phase and provides 13 ICPs for model assessment.

Reconstructions were conducted within APS under High alignment and Reference settings with positional data provided by the on-board camera GPS. Two models were created. First, a one based on nadir orientated images from 130 cameras generating a sparse cloud of 6.7 x 10^5 points. The second, by contrast, used images from both

cameras to create a convergent network comprising 292 cameras, yielding a more detailed sparse cloud of 1.1 x 10^6 points. The sparse cloud generated from each scenario was then initially optimized using the reduced set of GCPs (n = 14) described above. The resulting models were evaluated using the metrics described in section 3.7. In this instance the 13 GCPs left out of the geo-rectification stage within APS were used to provide ICPs.

Table 5.13 Error metrics derived from GCPs and ICPs for both Nadir and Convergent reconstructions of the town of Glenorchy. Convergent imagery consists of the same imagery within the Nadir only scan, with imagery from the additional camera set at 30^o aft-nadir.

Imagery	GCP MAE Pre-Optimization (m)	GCP MAE Post-Optimization (m)	LOOCV (m)	ICP MAE(m)	ICP MaxE (m)
Nadir	0.263	0.074	0.120	0.404	0.620
Convergent	0.105	0.075	0.084	0.177	0.290

As shown in Table 5.13 the convergent camera network generates a much higher quality reconstruction following the initial bundle adjustment with MAE (derived from positionally biased GCPs) of 0.263 m and 0.105 m for the nadir and convergent networks respectively. While this differential is mitigated following optimization using the GCPs as exterior constraints, it is notable that this effect has much less effect on the key lens model parameters for the convergent network compared to the nadir image set. For example, the optimization results in only a 0.28% change in the K1 parameter value for the convergent imagery, but a 1.76% change for the nadir image set (corresponding differences from the flume experiments were 0.24% for convergent and 1.54% for nadir). A LOOCV assessment, cycling the reduced GCPs as independent check points (n = 14), yields RMSEs of 0.120 m and 0.084 m for the nadir and convergent image sets respectively, again highlighting the quality of the convergent reconstruction.

Optimizing the model based on both the full (n = 27) and central (n = 14) GCPs, see Figure 5.13) provides a means to test the stability of the initial bundle adjustment. The impact of this is shown in the ICP MAE reported in Table 5.13 which are based on the 13 peripheral GCPs (i.e., 27-14) not used to optimize the model. In this case, the MAE falls from 0.404 m to 0.177 between nadir and convergent imagery respectively. In contrast,

the values recorded by the biased GCPs are 0.074 m and 0.075. This differential highlights the stability of the initial structure identified using the convergent imagery and the dependence of the nadir model on subsequent camera calibration across the entire area of interest.

The effect of the optimization process mirrors that seen in the flume experiments (shown in Figure 5.5). This is illustrated effectively by comparing DEMs of Glenorchy derived before and after optimization, shown here for both sets of GCPs (n_1 and n_2) in Figure 5.13. For both sets of GCPs the difference between the nadir only and convergent configurations is clear, with a bowl shaped deformation clearly evident by higher elevations around the perimeter of the optimized block. This deformation reaches up to 2 m within 500 m of the centre of the block and scaling this to the flume is an equivalent of 1.92 m, very similar to that observed in the experimental data. By comparison the convergent reconstructions appear much more stable and resilient to optimization with only minor deformation structures evident. This is particularly noticeable for the scenario based on optimization with the central (n_2) GCP set. In this case, the deformation for the nadir imagery is extreme with bowl –type distortions that become dominant in less than 200 m from the block centre, whereas the equivalent convergent model shows very little difference to that obtained with the whole GCP (n_1 = 27) set.

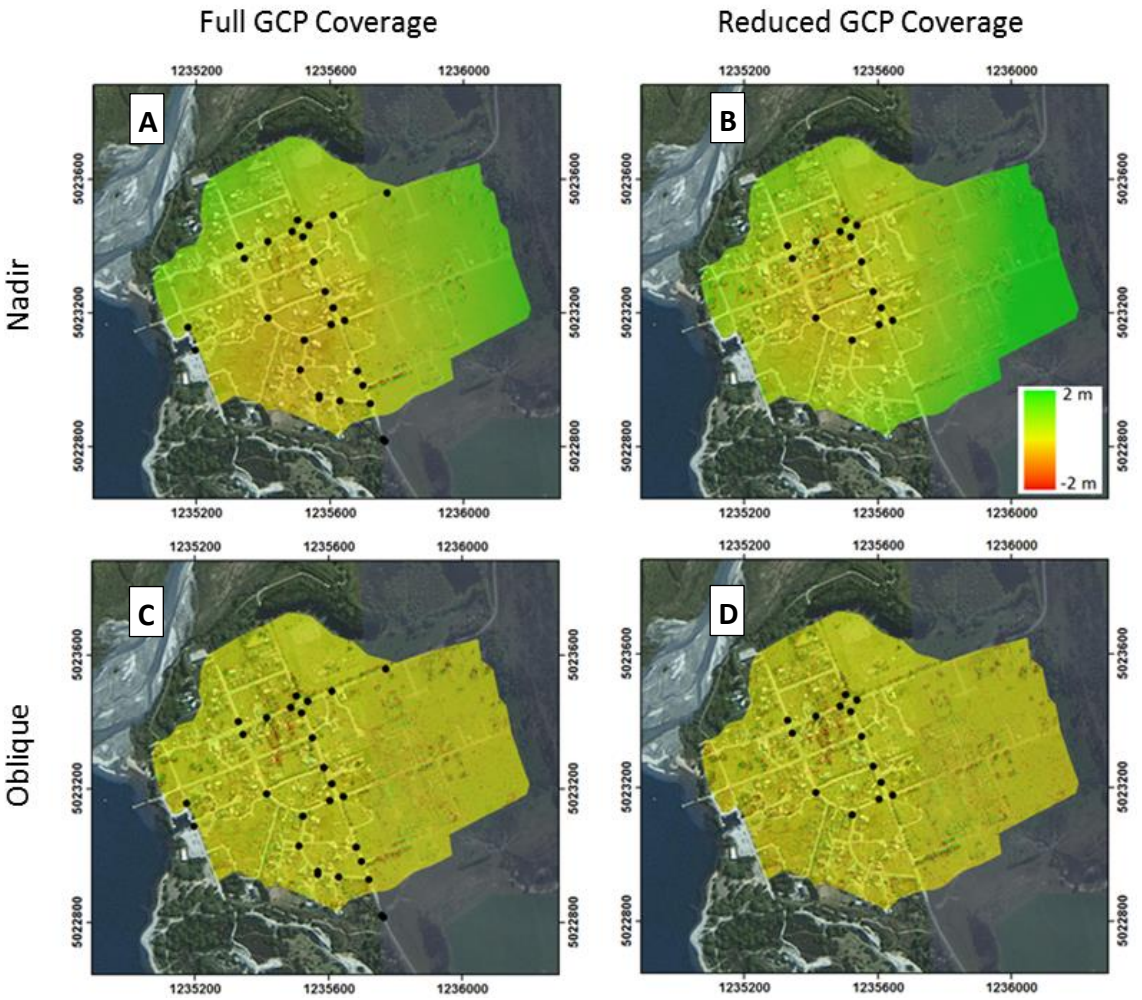

Figure 5.13 Optimization DoD (Post – Pre), displaying the lens correction applied to a) nadir imagery, full GCP, b) nadir imagery, reduced GCP, c) convergent imagery, full GCP and d) convergent imagery, reduced GCP. The convergent imagery exhibits minimal systematic change during the optimization phase in APS. *Projection: NZTM 2000*

5.4.2. Spatially Explicit Analysis of the Glenorchy DEM

The GCP based analysis described above is complemented by a more detailed spatially-explicit analysis of model quality that can be achieved through a comparison of the SfM derived models and the LiDAR survey acquired in 2011. Some care needs to be applied here given the time elapsed between the surveys (2011-2015), during which time there are considerable changes in terms of vegetation cover, disturbance of the channel bed and modifications to the urban plan and housing. However, even simple DoD analysis provides valuable insights into the differences between the nadir and convergent camera networks. Figure 5.14 shows two DoDs based on a comparison of a LiDAR digital

surface model (including vegetation and built structures) and the optimized nadir and convergent models described above, this time GCP n = 27. The analysis is based on the subtracting the SfM from the LiDAR (i.e., LiDAR – SfM) model, so positive values rendered here in shades of red, are where the LiDAR is higher and vice versa shown is shown blue. While the dominance of changes in the vegetation and built structures somewhat dominate, the convergent model is clearly well matched to the LiDAR data with over 95% of the surface lying with 0.25 m. By contrast, the nadir model is typically lower than the LiDAR and only 76% of the model lies within 0.25 m of the LiDAR dataset.

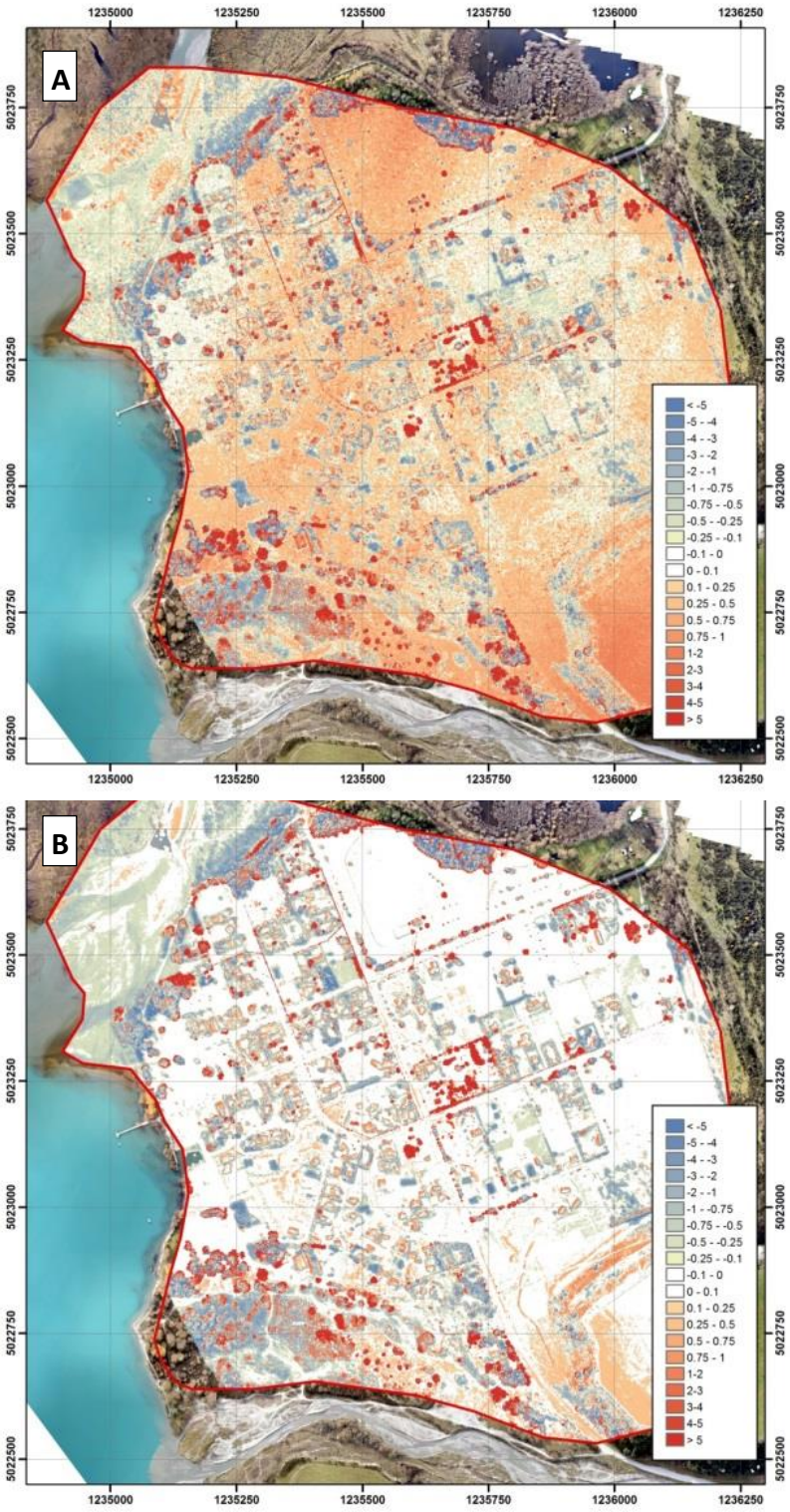

Figure 5.14 LiDAR-SfM DoD derived from A) Nadir only and B) Convergent imaging networks utilizing the full distribution of control points, n_1 = 27 GCP. *Projection: NZTM 2000*

5.4.3. Summary of Field Tests

The methodology to acquire convergent camera networks may differ depending upon the scale of the project and survey platform used. For example, relatively small scale studies conducted with UAS could achieve this using multiple flight lines with cameras orientated differently, flying at multiple heights, or circling an area of interest. However, for larger scale applications, where manned platforms represent the only time/cost effective platform, the ability (and cost) of implementing such complex flight plans is heavily constrained. The dual camera system used here does however, appear to generate an effective convergent network using simple parallel flight lines at constant height, greatly simplifying the acquisition process.

The experiments demonstrate that the use of the convergent network produces a more reliable initial bundle adjustment, without the use of GCPs as exterior constraints. When these are introduced to recalibrate the intrinsic camera parameters, the adjustment to the lens model is less significant, suggesting better a prior estimation of the key focal length and radial distortion parameters. The scale of these effects is also readily understood through the field trials, which are less sensitive to small survey errors that may confound the laboratory tests (i.e., many of the changes observed in these experiments are close the precision of the raw TLS and total station measurements).

5.5. Discussions

5.5.1. Impact of varying focal length and radial distortion

Results from the flume and field have clearly identified the benefits of a convergent camera network. Such image sets significantly reduce the necessity for posteriori lens calibration – a process which heavily reliant and biased by the distribution of GCPs. This reflects the improvement in *a prior* solution of the bundle adjustment using only inner constraints that is the first process in SfM photogrammetry (Cornelis et al., 2002).

In this study, it has been found that optimal results were obtained with cameras orientated at a 30^{o} angle. While reliable results were obtained with higher angles of separation, these effects will be magnified during the turning circle of the helicopter platform and high angles may ultimately incorporate additional noise through the loss of feature fidelity and matching performance. The selection of a 30^{o} angle is also

consistent with the findings of Wackrow and Chandler (2008) who found that two convergent cameras at 15^{o} to produce ideal results. Wackrow and Chandler (2008) suggested that convergent imagery acts to reduce systematic distortion by improving the self-calibration of the intrinsic camera parameters due to the enhanced depth estimation achievable with convergent imaging (see Chapter 2).

It is important to understand that the optimization process used in APS actually allows for compensation between the intrinsic (lens) and exterior (pose) orientation which are solved simultaneously. Close examination of this effect shows that adjustments to the estimated focal length are also compensated by changes to the exterior height of cameras. The interaction between focal length and pose estimation has been previously documented, with changes restricted to the domain of a translation in the Z dimension (Kumar and Hanson, 1994). It is important to consider the impact that focal length and its relation to exterior pose may have upon reconstruction geometry through both internal and external calibration and how this may interact with the presence of radial distortion.

This was achieved here, by defining a baseline reconstruction of the flume, *Scan A* (Scenario 3), which acts as a constant surface with baseline camera parameters. These baseline parameters were then altered individually in a simple perturbation analysis using a fixed camera calibration in which the K1 and Focal length were varied by 2% and 5% during the alignment phase. The resulting reconstruction was then rigidly transformed to the world coordinate system but not optimized. The output surface was then compared against the baseline reconstruction, to provide an insight into how the parameters affected the reconstruction directly. These results of this analysis are captured here as DoDs between the reference and parameter perturbation-surface presented in Figure 5.15. These results all relate to the convergent camera network.

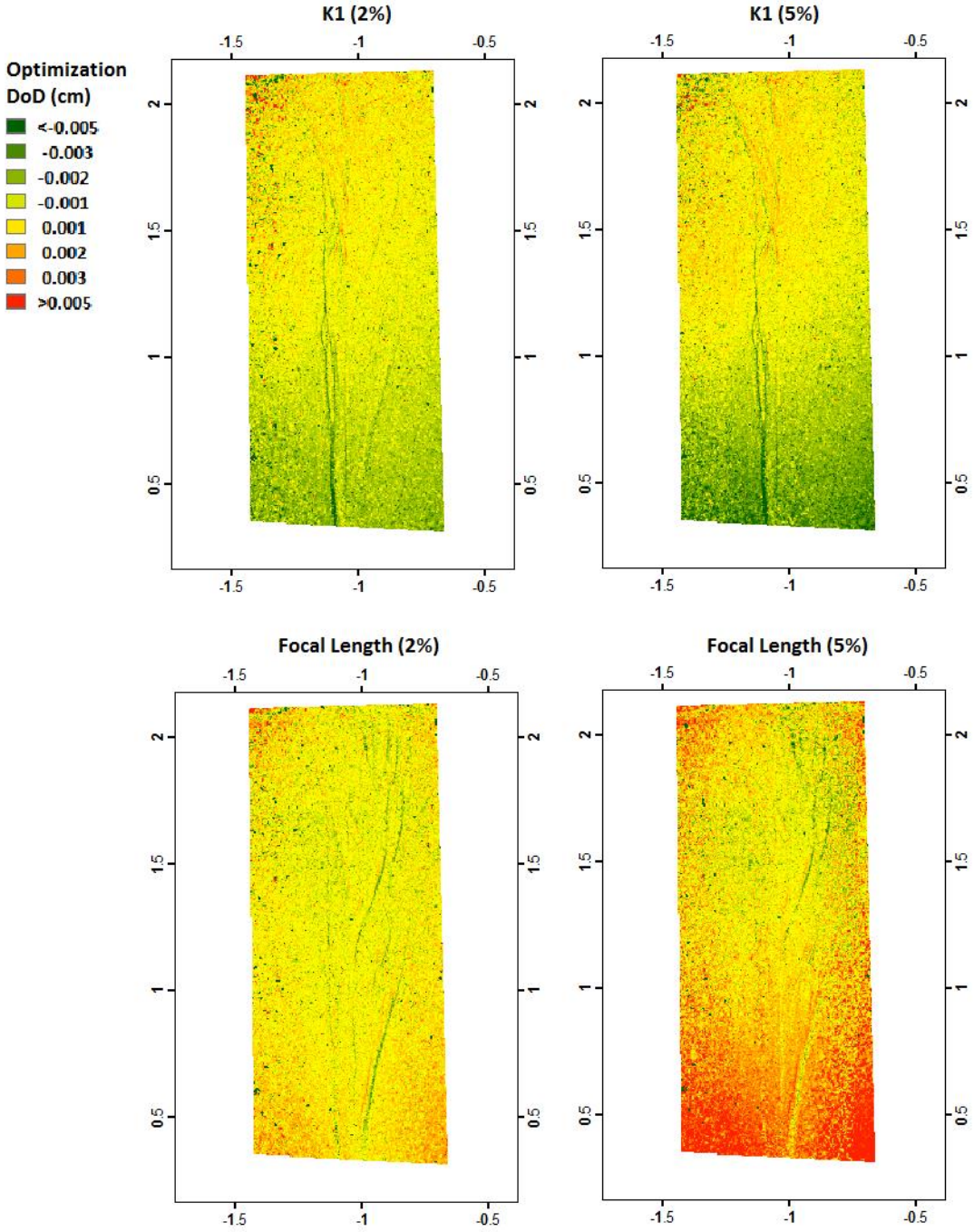

Figure 5.15 DoD generated from subtracting a reconstruction generated with a fixed calibration from a standard reconstruction. The changes observed related to the impact of maintaining a fixed calibration whilst perturbing the Focal Length and Radial distortion by varying degrees (2% and 5%).

This analysis reveals that the direction of change in the systematic error is opposed for each parameter. The change in magnitude of radial distortion perturbation is considerably less than that experienced by the focal length. In Figure 5.15 there is a significant degree of change for both altering the K1 parameter by 2% and 5%. However,

the focal length appears initially unperturbed at 2% with a dramatic change at 5%. This would imply that the impact of radial distortion parameter miscalibration is significant throughout, whereas the focal length only impacts past significant deviations. Similar results have been presented when considering the impact of both parameters on panoramic image reconstruction. Kang and Weiss (1997) noted that only past a significant threshold of misestimation did the focal length begin to result in systematic distortion, however, once reached this impact became the dominant source of distortion. Interestingly, the change in pose (in particular, height) as a result of changing focal length is proportional: 0.03 m for 2% and 0.07 m for a 5% adjustment respectively. In contrast the K1 parameter has no impact on the external estimation of camera positions with an average change in height of 0.0002 m observed consistently for both nadir and convergent reconstructions regardless of the magnitude of parameter adjustment K1. This would imply that the change between camera heights seen previously between nadir only and convergent networks is a result of the focal length. However, in the scenario presented here the calibration is fixed, whereas in reality the bundle adjustment simultaneously resolves these parameters. This complicates our ability to truly separate the individual impacts of each parameter as they are constantly adjusting simultaneously.

5.5.2. Impacts at the Field Scale

Optimization of the 2014 model gave rise to a series of irregular but oscillating corrections (shown by the post – pre-optimization DoD). This pattern is different from the simple bowl or dome-type distortions described by James and Robson (2014) which they suggest result principally from incorrect estimation of the radial distortion (K1) parameter. Such doming effects have typically been observed in studies of dense, well distributed ground control with careful flight planning to ensure even coverage of imagery (James and Robson, 2014). For wide-area or catchment-scale such as posed here, such idealized conditions are difficult to achieve due to the aforementioned restrictions on survey time and GCP placement. Inevitably, this results in an irregular distribution of ground control and poor initial external camera constraints, which in turn give rise to inaccurate estimation of focal length and pose.

The irregular pattern observed within the optimization phase is however assumed to be inherently present before the introduction of GCPs. The irregular shape of the area of interest combined with the impact of non-consistent pose errors throughout the survey is therefore responsible for this irregularity. The focal length has been seen to introduce a unique source of distortion beyond a significant threshold of calibration error and appears to be heavily linked to pose estimates. It is the impact of these erroneous pose estimates as initial constraints within the bundle adjustment upon focal length estimates, combined with the inherent radial distortion that results in an irregular pattern. By contrast, the dome observed in the study on the debris fan in Chapter 4 and within the flume are the results of no prior exterior constraints (introducing no error) over a well constrained regional block and as such radial distortion dominates.

It is important to recognize that this is inferred through fixed calibration experiments (Figure 5.15) that do not truly reflect the interaction between parameters during their simultaneous solving during the bundle adjustment. As such, APS remains a black box piece of software rendering it impossible to determine the exact impact of each parameter. Result here however would suggest that the way in which the camera calibration reconciles itself varies for each parameter, as erroneous pose estimates may manifest more actively in the compensation of the focal length.

5.5.3. Improving Computational Efficiency

As stated in Table 5.1, a side-effect of adding oblique imagery through a dual camera system is the inevitable doubling of an already large number of images. The experiments described here have sought to investigate whether the additional computation burden this creates can be mitigated by optimizing image density (by thinning the camera network) and by compressing the image format.

The flume experiments described here suggested that the image network could be thinned by as much as 66% for a convergent scenario (n = 18 vs. = 48) and compressed without substantial deterioration in model quality. At the same time, this reduction in complexity generated a 73% saving in computation time. The magnitude of potential time savings may permit the use of higher alignment settings for projects that are limited by the computational power available. Results from the flume also show that large initial improvements in nadir imagery reconstruction geometry can be made by

increasing image density at low frequencies to a threshold of 75% overlap. Past this point however, the returns in terms of model quality vs. image density diminish and serve only to add computation time.

For field surveys, optimizing the number of images required can be achieved through a reduction in capture frequency, through an increase in flight speed or by changes in the flight path. This is particularly relevant given the cost of helicopter flight time and the variation in lighting, particularly in areas of high topographic relief that occur during long surveys. Reductions in the required density of images could therefore be achieved by increasing flight speeds (and at increased altitudes providing an adequate ground pixel resolution is maintained by increasing camera resolution) which would have the benefit of reducing flying costs and minimizing lighting changes. The concomitant reduction in image numbers will also lead to greater computational efficiency. However, given the potential risks of over-thinning the image network and the much greater costs associated with remobilizing for another survey in the event of lost data, here a cautionary, conservative data acquisition is preferable. The effects of oversampling can then be optimized by selectively thinning the image network post-hoc. A compromise of a capture frequency at 1 Hz and an increased flight speed from 85 knots to 90 knots is therefore proposed. Additionally, the use of a full-frame SLR camera, such as the Nikon D750 will also increase the field of view enabling equivalent overlap from higher altitudes, while at the same time the 24 MP resolution of this camera will offset the effects of flying height on pixel ground spacing.

Compression of the captured imagery has also been shown to decrease computational requirements without significant deterioration in the model geometry. This chapter has assessed the application of a significant image compression in a lossy JPEG format. This has been seen to impact on the accuracy of keypoint identification and correlation, increasing noise in the scene structure. Despite this, the impact of capturing in a JPEG format will act to minimize logistic concerns experienced in the field when recording at high capture frequencies for large periods. The capture of RAW and lossless compression format files (TIFF), such as those acquired in 2014 (comprising >40 MB), are ultimately limited by the write speed of the recording format and transmission cables. A lossless JPEG compression, resulting in file sizes of ~ 15 MB will result in

reduced computational times, with less severe detrimental impact to the accuracy as experienced in this chapter. In 2014, trial and error in the field found that the maximum capture rate that could be achieved was only 2 seconds, greatly diminishing the density of images acquired, creating further problems down the processing chain.

5.6. Conclusions

This chapter has presented a series of scaled experiments designed to inform the survey requirements of subsequent surveys of the Dart River. In particular, the goal was to explore strategies to mitigate the systematic errors observed in the 2014 data modelling. The following conclusions can be drawn:

A. Convergent camera networks reduce systematic deformation arising in the SfM bundle adjustment. This is evidenced by comparison with both reference TLS data models and through comparison of before and after optimization surfaces. Incorporating imagery acquired from an angle of 30^{o} aft-nadir provides the optimum oblique angle to combine with vertical imagery. The use of convergent camera networks acquired with this viewing geometry relax the requirement for posterior optimization of the structure, a process which introduces a spatial bias through the distribution of the GCP network. Areas of sparse ground control (particularly the perimeter of the area of interest) are most affected. These results have been validated by a combination of both scaled and natural length scale experiments in the laboratory and field.
B. The source of the systematic error arises from misestimation of the intrinsic camera parameters (both the internal radial distortion and focal length) and through their compensating effects on the exterior camera orientation (pose). These effects are magnified by irregular GCP placement and irregular flight paths and in combination give rise to oscillating patterns of correction observed during the optimization.
C. Increasing image density up to a threshold (75% for nadir imagery) provides enhancements in the quality of reconstructions. However, beyond these thresholds there is little appreciable gain in quality and only cost in terms of increasing computation time. This result suggest scope to define optimal data

acquisition strategies based on shorted flight times (increase flight speeds). Reducing the image number also raises the possibility to use higher accuracy settings for the SfM bundle adjustment.

D. Derived surfaces exhibit relatively minor sensitivity to image compression. However, degraded imagery offers significant computation benefits, reducing processing time by an average of 19% for the experiments shown here.

E. The use of scaled laboratory experiments has been shown to have great potential to help define field scale survey practice. This approach offers cost savings and greater flexibility. However, the scale of experiments and the analysis of results is ultimately limited by instrumental measurement accuracy, and for the experiments shown here, many of the observable changes are close the limit of detectable change.

5.7. Implications for Future Research

Results from this study suggest a series of tangible modifications to the 2014 survey protocol. These are summarized in Table 5.14.

Table 5.14 Survey recommendations from the flume with corresponding advantages and potential issues and the subsequent implementation strategy for the 2015 Dart survey.

Survey Method	Advantages	Potential Issues	Implementation over catchment scales
Convergent Imagery	Improved geometry and a reduced reliance on ground control distribution	More imagery requiring additional computational time and increased sparse cloud noise	**Dual rig approach of 1) a nadir orientated and 2) a 30^0 aft-nadir orientated sensor**
Reducing Image Numbers	Improved computation efficiency with minimal degradation in results between 75% and 80%	Risk of under-exposure and extensive peripheral areas of minimal coverage	**Collect at maximum rate (1 Hz) at a flight speed of 90 knots and conduct removal post capture**
JPEG compression	Up to a 10% reduction in computational time and minimal degradation in results in a convergent network.	Loss of fidelity in high oblique regions and nadir only regions	**Capture in JPEG format but limit images used over a certain angle (i.e. when turning)**

The next stage of the process is to assess the impact of the survey modifications over a catchment scale characterized by un-even ground control and large imaging numbers. This will be in conjunction with results from Chapter 4, identifying the use of on-board camera GPS to improve pose estimations and reduce computational processing times.

6. Chapter 6: Topographic reconstruction of the Dart River, 2015: The systematic and localized impacts of using Structure-from-Motion with a convergent camera network

Abstract

Previously, this thesis has identified the benefits of a convergent camera network and improved pose estimates to minimize systematic errors and therefore, reliance on ground control. This chapter expands the SfM pipeline from a local block to a corridor mapping problem in order to reconstruct the Dart River study site in 2015. Following the survey recommendations outlined in Chapter 5, a total of 16,000 images were captured from two cameras operating concurrently (nadir and 30^{o} aft-nadir) from a helicopter at a flight height of 375 m travelling at a speed of 90 knots. These were used to generate a dense point cloud comprising 1.3 x 10^{9} points.

The convergent imaging geometry is shown to lead to an increase in noise within the sparse cloud (local surface SD of elevations ~0.4 m). However, analysis of the cloud following densification indicates an equivalency in the degree of noise between both nadir and convergent imagery. Much of the observed noise results from regions at the edge of the Field of View (FOV) which generate small numbers of poorly correlated features. These are however filtered out using the pairwise Semi Global Matching (SGM) densification algorithm. The incorporation of a camera GPS is shown to lead to unforeseen problems due to the high flight speed. This required significant manual correction to provide consistent co-ordinates to support the initial block bundle adjustment. Incorporating camera positional data does however represent an improvement on the heuristic approach undertaken within Chapter 4 and minimizes the extent and magnitude of pose inaccuracies.

The geometric quality of the 2015 Dart survey is assessed here using 87 Ground control targets as both constraints and check points and reveals MAE (n=20) of 0.23 m. Although evidence of systematic bias remains, it is greatly reduced in comparison to the 2014 Dart survey, contributing to the 50% reduction in check point error. These encouraging results provide a basis to use the 2015 survey as part of a DEM-of-Difference assessment to assess geomorphological change within the reach.

6.1. Introduction

6.1.1. **Challenges of catchment scale SfM**

Chapter 4 demonstrated the potential to derive wide-area (catchment-scale) terrain products through helicopter-supported SfM photogrammetry. However, a combination of poor camera network and ground control geometry, driven largely by logistical constraints in 2014 resulted in a sub-optimal survey that incorporated significant systematic errors. While the terrain products appear visually convincing, such errors limit the use of these data for change detection, a pre-requisite of which is reproducibility. The spatial manifestation of these errors is a result of compensation between internal and external calibration parameters during the SfM bundle adjustment. These effects are severe in areas of sparse control, such as around the perimeter of the survey. Controlled, scaled laboratory experiments and field validation, has subsequently demonstrated the benefits of convergent camera networks and *a priori* pose estimation (i.e., through a camera GPS) to minimize errors and relax the reliance on ground control.

The revisions to the survey strategy proposed Chapter 5 are designed to mitigate the logistical challenges associated with applying SfM at large fluvial system-scales and in hostile environments. The rationale behind these adaptations is detailed in Table 6.1.

Table 6.1 Approaches chosen to address the majority of logistical and environmental challenges posed by catchment scale surveys, as initially listed in Chapter 4, Table 4.1

Survey Strategy	Environmental/ Logistical Challenge	Advantages for catchment scale surveys	Potential Issues of application
Convergent Imagery	• *Large Survey Extent* • *Labile Environment* • *Complex topography*	Reduced impact of radial distortion and improved camera height/focal length estimates. Less reliance on unreliable GCP network restricted by hostile remote environment. Potential to reduce human cost by reduced ground control placement. Secondary angle increase coverage of marginal regions where ground control is often placed and in regions limited by flight lines.	Increase in cloud noise through high oblique shots with increased pixel distortions; Most cost effective implementation results in a doubling of image count
Camera GPS	• *Large Survey Extent* • *Labile Environment* • *Computational Resources*	Improved computational efficiency during the bundle adjustment and initial estimation of exterior pose, beneficial in regions of limited external control.	Reliability of GPS mounted beneath a helicopter and potential latency issues when used at fast flight speeds.
Imagery format	• *Large Survey Extent* • *Financial Constraints* • *Computational Resources*	Reduced processing times in the bundle adjustment and the ability to write to file quicker in the field enabling increased capture frequency and flight speeds.	Degradation in pixel quality in homogenous areas where little contrast exists.

6.1.2. **Convergent Camera Network**

The role of convergent camera networks in reducing systematic errors is covered in detail in section 2.5. Briefly, convergent camera networks increase geometric consistency by improving the estimation of depth modelling of radial distortion (Wackrow and Chandler, 2008; Wackrow et al., 2011; James and Robson, 2014). The application of a convergent network within the town of Glenorchy (section 5.4) reduced independent control point (ICP) errors (MAE) by over 50% and maximum observed errors by 65% compared with a nadir orientated survey. Furthermore, the application of a convergent network has been seen to minimize the degree of change during the optimization phase within APS and thereby reducing reliance on ground control; a significant limitation in this environment. Acquiring convergent photography using a dual sensor system also presents the opportunity to minimize flight time, although exacerbates the computational burden by doubling the size of the image set. Such an

approach will however provide greater coverage in peripheral locations where placement of ground control is limited and in inaccessible regions like the gorge. Finally, the increase in image density will also support enhanced transformation and subsequent calibration of the sparse cloud due to the additional redundancy in the identification of GCPs (Agisoft, 2017).

Whilst there are clear benefits to using a convergent imaging network, evidence from Chapter 5 also suggested that oblique imagery may contribute to localized errors and surface noise. The Glenorchy trial experiments did however indicate that these anomalies were greatly reduced during densification of the model. This is assumed to be a result of the densification strategy used in APS. Unlike some other approaches to depth mapping which explicitly incorporate the sparse cloud, the SGM algorithm uses the resolved camera geometry independently. First to identify stereo-pairs to resolve depth at the pixel resolution by intersection and then average the results of multi-stereo pairs. However, as APS is weakly documented software, it is important to assess the performance of this process in areas of increasingly complex topography. Importantly, it is necessary to understand the extent to which such noise, potentially indicating spurious tiepoints, propagates into the estimation of the camera geometry and therefore influences the independent dense alignment.

6.1.3. **Camera GPS**

The use of an on-board camera GPS will improve processing time and accuracy of the SfM bundle adjustment. Processing the bundle adjustment under the Disabled alignment setting (see section 3.4) in APS involves an image by image multi-view analysis which results in a factorial relationship between image numbers and match comparisons. Reference alignment settings (which require information on camera position), by contrast, use pre-identified pairs to constrain comparisons regionally. This constrains the parameter space of the bundle adjustment, improving both the processing speed and reliability. The Reference setting is therefore vital for large scale reconstructions involving >10,000 images under high image quality settings (i.e., native image fidelity).

The difficulties encountered in reconstructing the Dart in one photogrammetric block without camera GPS were documented in Chapter 4. This necessitated the development

of a heuristic workflow to infer positional camera information in the Gorge reach using low quality data sources such as Google Earth. The propagation of errors associated with these low quality exterior constraints was responsible for inaccurate pose estimation that developed into systematic errors across the whole model. The use of an on-board GPS receiver should improve pose estimates and reduce the impact of incorrect focal length estimation which was identified as a main component of the pattern of change observed during the optimization process.

The quality of camera positioning by GPS is strongly determined by the type of receiver used. Low-cost, directly integrated sensors such the Solmeta Pro 2 used here, are typically code-only systems with accuracies quoted of the order 10 m. These estimates are however associated with stationary measurements, but in this case are needed on a camera travelling at 90 knots and attached to the helicopter skid which will restrict the field of view. There is a significant risk that GPS latency in the GPS solution and subsequent undocumented positioning averaging could introduce significant errors in the estimation of 3D position. Moreover, as flight speed will vary over the flight path (slowing during turning at the end of transects), there is further likelihood that such errors will vary through time. Studies using UAS over smaller photogrammetric blocks are typically flown at lower speeds (10-25 knots), with more easily programmed flight paths and are therefore subject less to these concerns.

6.1.4. **Aerial Photography**

The 2015 survey described below was conducted at an increased flight speed capturing imagery in JPEG format following the successful laboratory tests. The increased flight speed was possible due to the greater redundancy in the image set given the use of two cameras. Additionally, the use of image compression facilitated stable recording of images on the on-board data cards at higher frequency (1 s) than used in 2014 (2 s). The faster write speed also enabled concurrent writing of data to a laptop within the helicopter – which provided additional backup and the opportunity to verify successful capture and refine exposure setting. These changes reduce the total flight time, maximizing financial resources, but also limiting the impact of changing illumination through the survey. The use of image compression and increased flight speeds may compromise the quality of keypoint identified during the alignment phase as pixel

sharpness is reduced. This is particularly relevant for identifying features in homogenous regions and at the FOV of high oblique images, potentially resulting in a change in keypoint scale and increased surface noise in the sparse cloud.

The additional redundancy allows for the possibility to remove extensive imaging numbers from overly captured regions post-capture. The combination of both increased flight speeds and reduced image file size enable sufficient data coverage across all regions with reduced processing times, flight costs and deteriorations in survey quality associated with changing light conditions. In order to partially offset these effects, a higher quality, full frame SLR camera was used, the Nikon D750. This creates an increase in the field of view by approximately 34%, therefore imaging a larger ground area than observed in 2014, or permitting an increase in flying height, simplifying the flight path, with the reduction in ground space pixel dimensions offset by the greater resolution of the sensor CCD (24 vs. 12 MP).

6.1.5. **Chapter Objectives**

Chapter 5 provides a trial of the survey strategy described above over the township of Glenorchy. This experiment demonstrated marked improvements in model geometry, although it should be noted that these results were obtained over a small photogrammetric block with dense ground control distribution and high image overlap. Up scaling the refined survey to the full 35 km reach presents a more challenging task, still constrained by the low density and irregularly placed ground control and complex scene geometry. This chapter aims to provide a detailed assessment of the new survey protocol and evaluate the fitness of the derived terrain products for further geomorphological research. This will be achieved through the following research objectives:

Objective 1: Evaluate the impact of the 2015 survey refinements upon random error through analysing noise within the sparse cloud and the impacts this may have upon dense reconstructions.

Objective 2: Implement the SfM workflow refined in previous chapters to generate a 3D reconstruction of a 35km reach of the Dart River.

Objective 3: Quantify the magnitude and structure of uncertainties in the resulting DEM.

6.2. The 2015 SfM reconstruction of the Dart

6.2.1. Image Acquisition

The image acquisition strategy for the 2015 survey is described in detail in Chapter 2. Briefly, a convergent network of images was captured using two Nikon D750 cameras orientated at nadir and 30^{o} aft-nadir mounted concurrently to individual rigid mountings attached to a helicopter skid.

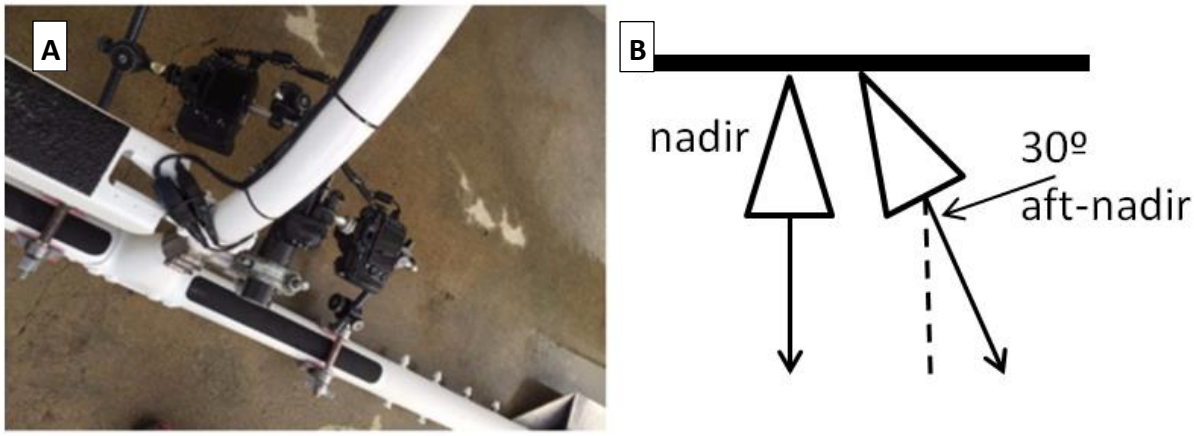

Figure 6.1 A) Implementation of the convergent camera network upon the skid plate, to provide aft-nadir and nadir imagery and B) schematic of the angle.

Images were acquired with a Nikon D750 camera, with a fixed 28 mm lens with focus set at infinity and exposure settings f/6.3, 1/1000s. Images were recorded in a lossless JPEG format with an average size per image of 15 MB. The flight plan was structured into a series of sections approximately 5 km in length and flown at 400 m above ground level (AGL) at a velocity of 90 knots.

The flight lines were constrained by the valley geometry but broadly reflected the same pattern used in 2014 (Figure 6.2). Images were acquired in quasi-parallel flight lines, planned to provide both sufficient side and endlap and were implemented by an experienced pilot used to the demanding flying conditions. The gorge section comprised three transects, offering six views which provided sufficient coverage to reconstruct the region unlike 2014. Images were triggered automatically at a frequency of 1 Hz resulting in a network of >16000 images with an average endlap and side-lap of 90%. This is, however, an over-estimate as images captured during turns and associated with very

oblique angles were subsequently discarded. Flight time was scheduled around midday to coincide with the solar zenith and to minimize the impact of oblique sun angles. Each camera was attached to an on-board GPS geo-tagging imagery which was subsequently written to SD cards on the camera and relayed as back-up to two laptops on board the helicopter.

In line with the procedure detailed in section 3.4 and undertaken in Chapter 4, the image set was reduced to 10,726 by manually removing images of poor quality, over obliquity and those comprising mainly of vegetated or wetted surfaces. Imagery was loaded within APS and an alignment was processed on a Dell Optiplex Dual Core I7 with 16 GB RAM and 1 GB NVIDIA graphics. The high alignment setting was chosen alongside the reference setting to make use of the camera GPS data available and reduce the computational time. The alignment resulted in the reconstruction of an initial sparse point cloud comprising 44 million points. The initial sparse cloud was however prone to excessive noise about the surface, regions of incoherent geometry and significant camera errors (the difference between the initial *a priori* camera position and posterior position). These issues are addressed in the following sections (6.2.2 and 6.2.3) and precede the construction of the final point cloud for the 2015 Dart survey.

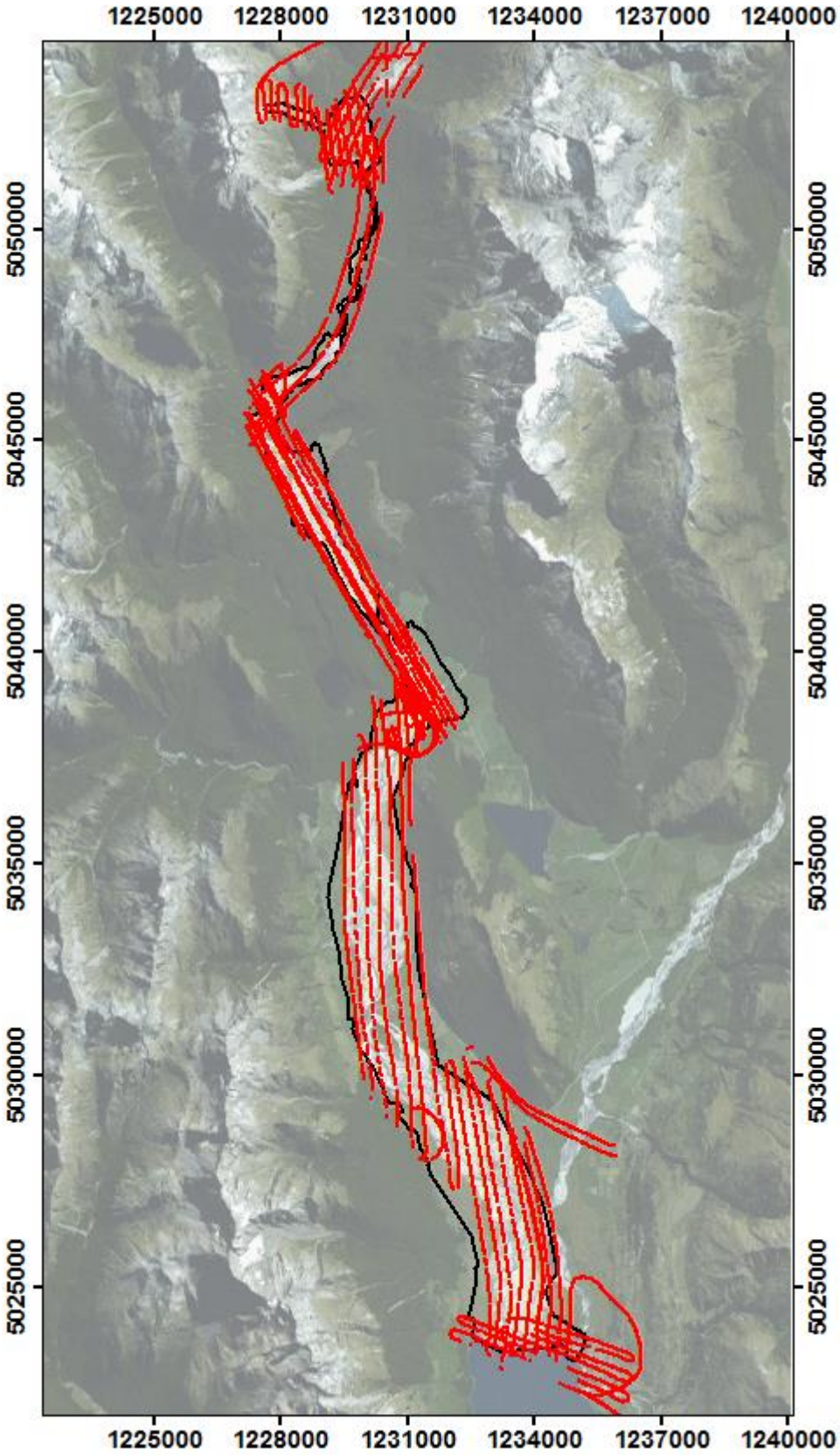

Figure 6.2 Image capture location by GPS reading for the 2015 survey. Each point represents the capture of two images, orientated nadir and 30^0 off nadir. *Projection: NZTM 2000*

6.2.2. Noise within the sparse cloud

6.2.2.1. Variance of sparse cloud and dense alignment

The experiments documented in Chapter 5 indicate additional surface noise associated with convergent camera geometry. It is important to assess the origin of this noise and to determine if this: i) may be reduced by further adapting the survey strategy; and ii) whether this noise propagates into the final dense cloud by contributing significantly to errors in the solution of the interior and exterior camera geometry.

Identifying the source of sparse cloud noise is complicated as convergent imagery contributes not only multiple viewing angles, but though the dual camera system, additional images and so reduces the effective baseline distance. This latter effect has previously been highlighted as a source of noise in APS based reconstructions (Marcis, 2013). Reduced baseline distances have been shown to increase reconstruction uncertainty as the 3D position of point is based on the aggregation of results from a wider set of stereo-pairs, incorporating a large range of baselines. The effect of baseline distance on depth estimation can be visualized by considering a simple stereo pairs as shown in Figure 6.3.

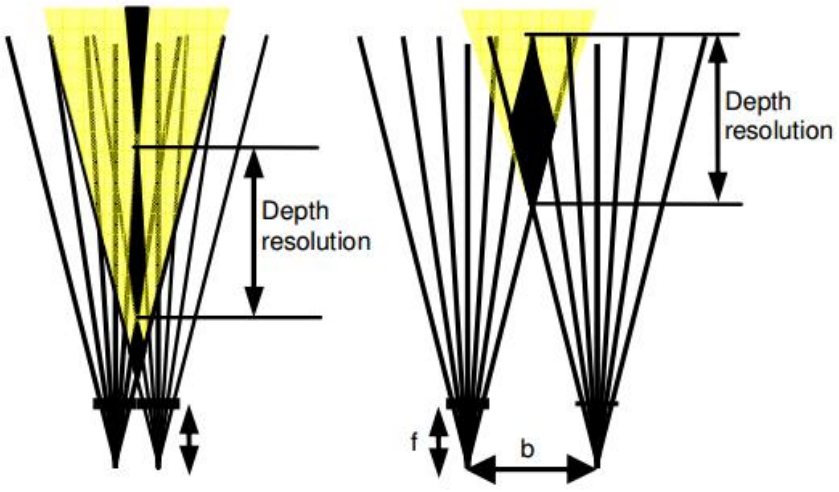

Figure 6.3 Schematic representation of changing baseline upon depth estimation, where f is focal length, b is the baseline distance. As the baseline distance is reduced, the potential region of location within the Z orientation increases (source from, Kyto et al., 2011; p. 3).

As the baseline distance is reduced it leads to an increasingly parallel angle between sensors, which minimizes the error in the x-y plane, but increases the potential error in the depth estimation. After Kyto et al. (2011) the theoretical depth resolution dZ_c at a distance Z may be calculated as:

$$dZ_c = \frac{Z^2}{fb} dp_x \quad [6.1]$$

where f is focal length, b is baseline distance and dp_x is disparity accuracy, which relates to the matching accuracy of the correspondence algorithm chosen. In this instance the depth resolution scales linearly with the baseline distance.

To identify the source of noise, a series of tests were conducted here by building a small photogrammetric block under a range of different scenarios (Figure 6.4). These included both convergent and nadir imagery as well as the full image density and halving the image numbers to increase the baseline distances in the camera network. As well as the basic analysis of check points errors across the whole block, three 'flat' gravel bars of approximately 100 m^2 (see Figure 6.4) were used to provide detailed statistical analysis of the resulting noise based upon both the raw sparse and dense cloud. These bars (denoted here after as U1S, U2S, and U3S) were selected due to their uniform, dry character and so are devoid of strong shadows. These areas were referenced by four GCPs, selected in a deliberately misleading configuration to promote further geometric deviations to contrast the impact of systematic vs. random error. This analysis was conducted via a LOOCV to determine the geometric robustness of each scenario.

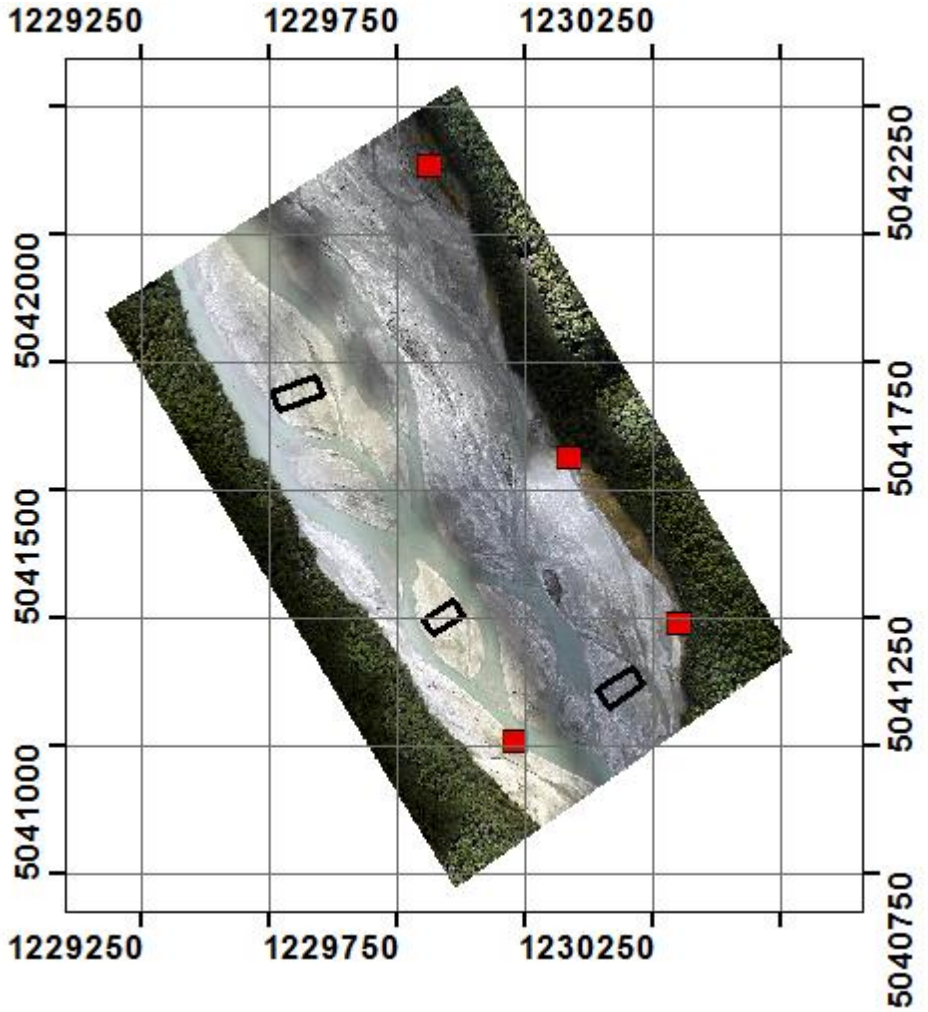

Figure 6.4 Orthophoto of the study region for the densification algorithm, with black rectangles representing the location of each bar sample and red markers the GCP distribution. This region was chosen specifically as channel reconstruction is relatively un-noisy for this photogrammetric block. *Projection: NZTM 2000*

*6.2.2.2. **Sparse Cloud Alignment***

Seven different reconstruction scenarios were developed to examine the effects described above alongside a reconstruction obeying the standard convergent workflow. This set includes both convergent and nadir only images, high and low accuracy alignment settings, and full and half density image sets. The summary of each model along with quality assurance statistics, derived using LOOCV based on the four local GCPs is presented in Table 6.2. DEMs of the resulting surfaces are also shown in Figure 6.5. The mean keypoint size and LOOCV results will be referred back to throughout this section as metrics of geometric accuracy.

Table 6.2 Combinations used to test the impact of alignment setting, sensor geometry and baseline distances upon the bar section, complete with reconstruction metrics. For example, Scenario: Low Full consists of full convergent imagery used for survey extent with standard baseline distances under the low alignment setting within APS.

Scenario	Alignment Setting	Imagery Used	Baseline Distances	Tiepoints	Reprojection Error (pix)	Mean Keypoint Size (pix)	LOOCV (m)
Low Full	Low	Full	Standard	7.1×10^5	2.75	15.32	0.77
Low Nadir	Low	Nadir	Standard	4.5×10^4	2.62	15.83	0.66
Low Oblique	Low	Oblique	Standard	8.6×10^4	2.48	15.36	0.95
Low Halved	Low	Full	Doubled	5.6×10^4	2.65	15.40	0.93
High Full	High	Full	Standard	7.1×10^5	0.89	3.80	0.41
High Nadir	High	Nadir	Standard	3.5×10^5	0.75	3.59	0.55
High Oblique	High	Oblique	Standard	4.8×10^4	1.07	4.70	0.27
High Halved	High	Full	Doubled	4.1×10^5	0.80	3.77	0.89

In Figure 6.5, the high alignment scans are shown on the right and exhibit a significant reduction in noise for both wet channels and bar surfaces. The addition of oblique photography (i.e., a convergent camera network) is, however, clearly associated with an increase in the surface noise, most evident in the channels. All model scenarios based on the Low quality alignment setting have particularly poor representation of the wet channels, with significant surface noise. Indeed, the high variability of elevation values in the wet channels could even be used to classify these areas effectively based on the local variance. Nadir orientated imagery alone appears to be less susceptible to noise, although the weaker performance using convergent photography can be offset using the High alignment setting, even with the full (standard baseline) image set. It appears from this simple analysis that the key limiting factor is the alignment quality setting. This parameter controls the resampling of the initial imagery and thus the size of the keypoints that can be identified. This is clearly show in Table 6.2 which shows typical keypoint sizes of 4 pixels for the high alignment setting versus 15 for the low quality setting.

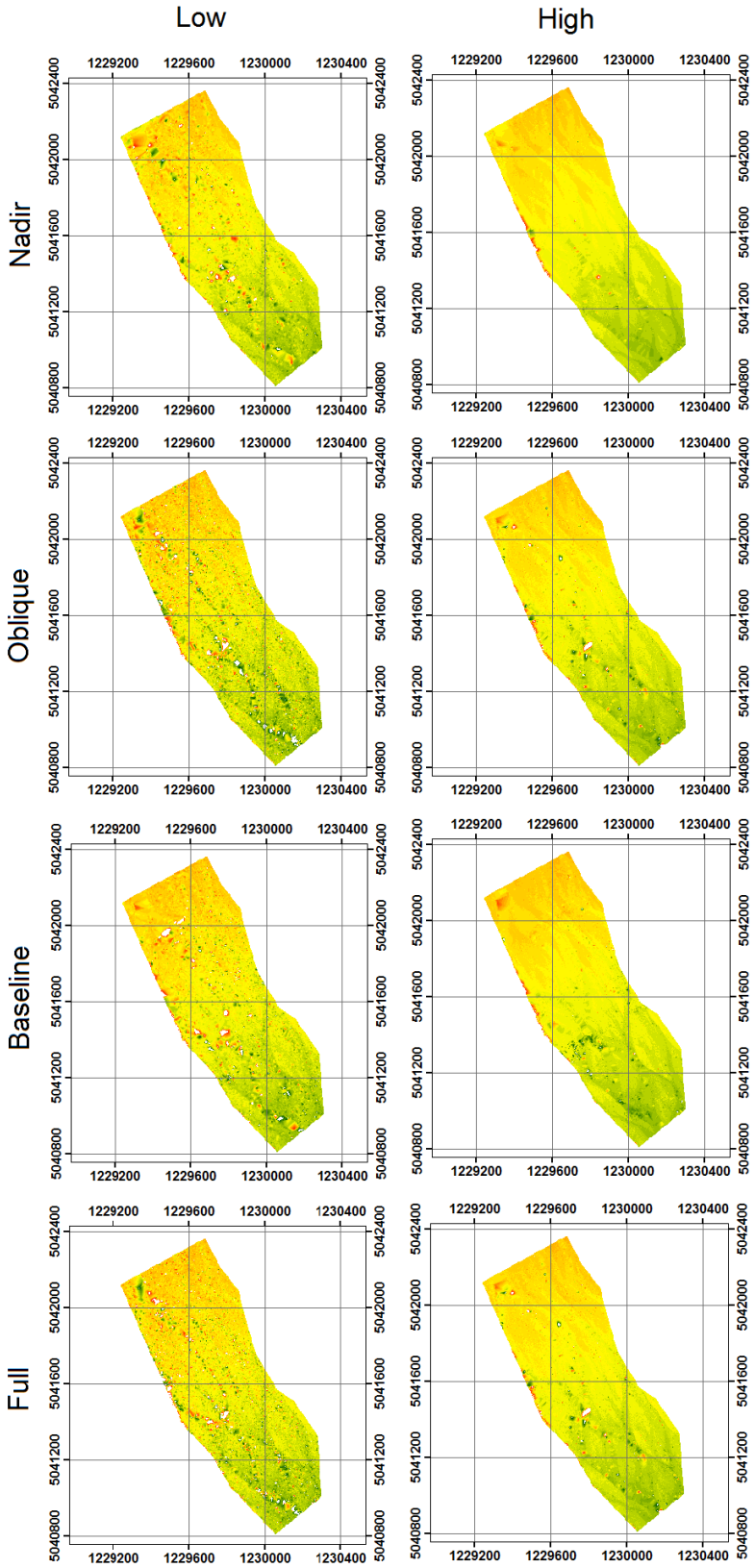

Figure 6.5 DEMs generated from sparse cloud for each tested scenario, where High and Low reflect APS alignment settings, Nadir represent nadir only imagery, Convergent represent convergent only imagery, Baseline represents halved image numbers to increase the baseline distance and Full represents the full convergent imagery (nadir and oblique). The DEM format is able to highlight regions of noise as they show up as rapid localized regions of variance. *Projection: NZTM 2000*

This analysis can be refined by examining the local variability of elevations in the three bars (U1S, U2S and U3S) described above. This is captured by the standard deviation (SD), interquartile range and mean absolute deviation about the mean of elevations of each test bar surface (Table 6.3). Given that the surfaces have minor surface trend, these variance measurements capture the apparent 'roughness', the dominant contribution to which is noise rather than particle scale rugosity.

Table 6.3 Statistics for the spread of the sparse cloud for each bar section surface and resulting averages. Mean absolute deviation represents the mean displacement about a calculated mean value for the surface, whilst IQR is the Inter Quartile Range.

Bar Section	Scan	SD (m)	Mean Absolute Deviation (m)	IQR (m)
U1S	Low Full	3.07	1.34	1.07
	Low Nadir	0.89	0.53	0.60
	Low Oblique	2.43	1.31	1.21
	Low Halved	1.28	0.76	0.89
	High Full	0.41	0.28	0.38
	High Nadir	0.28	0.21	0.33
	High Oblique	0.45	0.33	0.49
	High Halved	0.29	0.22	0.35
U2S	Low Full	5.28	1.70	1.02
	Low Nadir	3.73	1.37	0.83
	Low Oblique	2.37	1.24	1.17
	Low Halved	3.76	1.60	0.78
	High Full	1.23	0.52	0.37
	High Nadir	0.45	0.23	0.27
	High Oblique	0.44	0.33	0.50
	High Halved	0.91	0.42	0.41
U3S	Low Full	3.34	1.43	0.99
	Low Nadir	1.83	0.83	0.63
	Low Oblique	2.72	1.39	1.20
	Low Halved	2.04	0.85	0.62
	High Full	0.52	0.31	0.41
	High Nadir	0.32	0.23	0.36
	High Oblique	0.57	0.37	0.48
	High Halved	0.58	0.28	0.33
Average	**Low Full**	**3.90**	**1.49**	**1.02**
	Low Nadir	**2.15**	**0.91**	**0.69**
	Low Oblique	**2.51**	**1.31**	**1.19**
	Low Halved	**2.36**	**1.07**	**0.76**
	High Full	**0.72**	**0.37**	**0.39**
	High Nadir	**0.35**	**0.22**	**0.32**
	High Oblique	**0.49**	**0.34**	**0.49**
	High Halved	**0.59**	**0.31**	**0.36**

Reconstructions consisting of both nadir and oblique imagery (full imagery) exhibit the highest level of surface noise. The full convergent imagery shows a clear a reduction in the interquartile range in comparison to SD, which suggests the presence of significant outliers. The Maximum SD observed is a value of 5.28 m for bar section U2S under low alignment settings, although significant values of 3.07 and 3.34 are recorded for U1S and U3S respectively. Beyond the elevated noise associated with the full image set, there is little consistency amongst the remaining scan scenarios. However, all scans are improved significantly with the use of High alignments settings. The following analysis considers the average values derived from the bars taken together.

The implementation of greater baseline distances, by removing alternate imaging locations has shown a reduction of noise in the sparse cloud – by 60% and 30% for the low and high alignment, respectively. It is not axiomatic that this is sole cause however. Further analysis conducted within APS was able to determine the tiepoints responsible for generating noise within the cloud. Typically these are matched in peripheral regions of the oblique imagery at the very edge of the Field Of View (FOV).

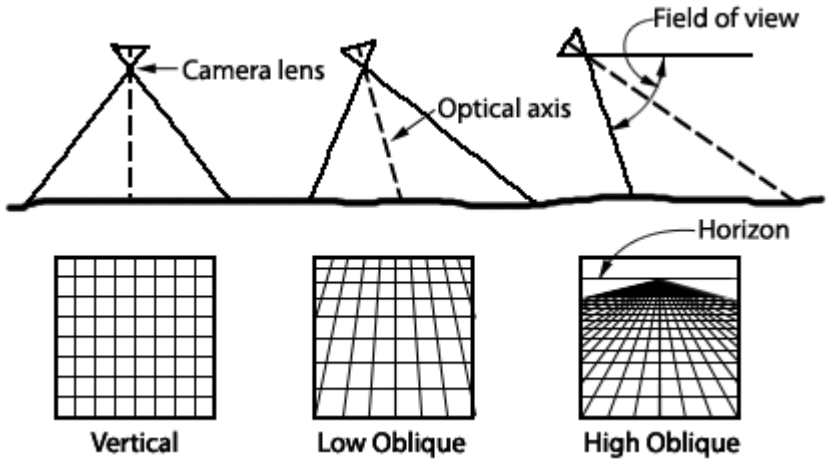

Figure 6.6 Schematic example of photo orientation on relative pixel values distributed across the FOV. High Oblique angles result in a loss of pixel fidelity at increasing distances from the sensor (figure source, Wolf and DeWitt, 2000).

Traditional stereo photogrammetry is prone to inaccuracies at the edge of the FOV with uncertainties characterized as a function of point location and change in viewpoint (Yang et al., 2010). This is exacerbated under oblique imaging conditions (Figure 6.6) as pixels at the edge of the field of view can be seen to be less reliable and less sharp resulting in

both erroneous and error strewn matches. The impact of a loss of sharpness corresponds to the point resolution and can be seen through the exaggerated decline using low alignment settings. This is supported further when considering the average keypoint size across the scans (Table 6.2), with an average value of 3.95 pix for high alignments and 15.45 pix for low alignments. Under the High alignment setting, oblique only imagery prone to increased edge of FOV related errors leads to a mean keypoint size of 4.70, significantly larger than the nadir only (3.59), full imagery (3.80) and halved imagery scenarios (3.77). Additional scans for both the Low and High alignment settings with the full imagery set were undertaken with the edges of the oblique imagery masked before reconstruction. Subsequently a surface SD value of 5.25 and 0.08 were recorded respectively indicating a significant reduction in noise when removing pixels undergoing significant warping at the edge of the FOV. Such an approach remains impractical reach-wide however as it impossible to know if a universal mask would not eliminate unique and valuable keypoints.

To identify the impact upon systematic error, the model is more appropriately tested via a LOOCV approach across the GCPs used to reference this subset of the Dart (Table 6.2). The LOOCV analysis highlights that halving the imagery and thus increasing baseline distances, consistently leads to a reduction in geometric accuracy, potentially resulting from the low image overlap or loss of imagery containing GCPs. While High alignment settings with nadir orientated imagery incorporate less surface noise, the surface geometry is still subject to a LOOCV of 0.55 m in contrast to just 0.27 m for oblique only and 0.41 m for the full imagery scenarios. The GCP analysis is able to discern that either oblique only or the full survey convergent approach provide the most geometrically accurate surfaces, necessitating the uptake of convergent imagery to minimize systematic errors, regardless of the noise within the sparse cloud.

*6.2.2.3. **Dense Cloud Processing***

The next stage in the SfM process is to apply the densification algorithm to assess whether variance within the sparse cloud propagates down the processing pipeline. The SGM algorithm used for densification comprises of two processes: i) a two tier penalty system is used to differentiate between actual surface change and outliers when producing disparity maps and; ii) a user prescribed confidence interval is set to control

the averaging of point estimates and produce a fused depth map from the stereo matches (Hirschmuller, 2008). These effects relate to the choice of filtering method employed in APS (see Literature Review, section 3.4). The analysis of the surface variance of each test bar following densification is shown in Table 6.4:

Table 6.4 Statistics for the spread of the dense cloud for each bar section surface and resulting averages. Mean absolute deviation represents the mean displacement about a calculated mean value for the surface. Aggressive filtering settings have been used within APS to derive the dense cloud.

Bar Section	Scan	SD (m)	Mean Absolute Deviation (m)	IQR (m)
U1D	Low Full	0.20	0.15	0.23
	Low Nadir	0.19	0.15	0.24
	Low Oblique	0.20	0.15	0.23
	Low Halved	0.21	0.16	0.23
	High Full	0.20	0.15	0.24
	High Nadir	0.19	0.15	0.24
	High Oblique	0.20	0.16	0.24
	High Halved	0.19	0.15	0.23
U2D	Low Full	0.13	0.10	0.17
	Low Nadir	0.12	0.09	0.17
	Low Oblique	0.14	0.11	0.17
	Low Halved	0.13	0.10	0.17
	High Full	0.13	0.10	0.17
	High Nadir	0.11	0.09	0.15
	High Oblique	0.14	0.11	0.18
	High Halved	0.14	0.11	0.17
U3D	Low Full	0.17	0.14	0.24
	Low Nadir	0.17	0.14	0.23
	Low Oblique	0.17	0.14	0.24
	Low Halved	0.17	0.13	0.24
	High Full	0.17	0.13	0.23
	High Nadir	0.17	0.14	0.24
	High Oblique	0.16	0.13	0.21
	High Halved	0.17	0.14	0.24
Average	**Low Full**	**0.17**	**0.13**	**0.22**
	Low Nadir	**0.16**	**0.13**	**0.21**
	Low Oblique	**0.17**	**0.13**	**0.21**
	Low Halved	**0.17**	**0.13**	**0.21**
	High Full	**0.17**	**0.13**	**0.21**
	High Nadir	**0.16**	**0.13**	**0.21**
	High Oblique	**0.17**	**0.13**	**0.21**
	High Halved	**0.17**	**0.13**	**0.21**

Variation between individual bars remains, however there is a significant universal reduction in SD, Mean Absolute Deviation and IQR in comparison to the sparse clouds.

Mean Absolute Deviation and the IRQ varies by just 10% across all scenarios and with similar results for the SD. There is therefore no major discernible trend or different between the reconstruction scenarios. Comparing the variance against that observed in the sparse cloud, it appears that densification reduces the SD by 95% under the high alignment setting and some 500 times that for the low alignment. This is principally a result of the pairwise approach. In this only the camera geometry (intrinsic parameters and pose) is carried forward, while the sparse cloud is only used to determine and seed the analysis of stereo-pairs. This enables the generation of comparable point densities regardless of the alignment quality setting. Variations may exist however between scenarios as differing sparse clouds may result in differing seeding orders for the implementation of the SGM. Once more, the nadir cloud under high alignment settings exhibits the least surface variance, however, the magnitude of variance across all scenarios is sufficient for subsequent geomorphological analysis. The improvements in surface noise are clearly evident in the DEMs generated for both the Nadir and Full scans in Figure 6.7, especially when compared to Figure 6.5.

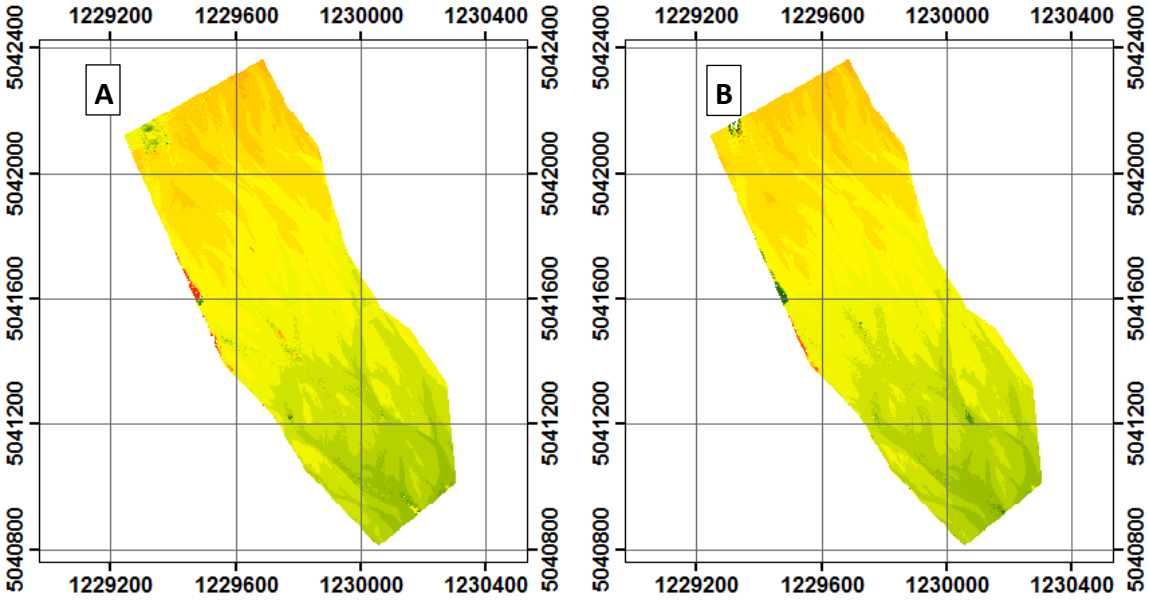

Figure 6.7 Dense DEMs generated for A) Nadir only imagery, B) Full Convergent imagery. Bar noise is seen to be equivalent with a few variations in water surface. *Projection: NZTM 2000*

*6.2.2.4. **Oblique Only Imagery***

The results described above from the LOOCV analysis along with findings in Chapter 5 highlight the importance of incorporating oblique imagery with nadir orientated views

and the impact of systematic errors in locations with minimal ground control. The high alignment settings represent a step change in quality of the sparse cloud, seemingly in relation to the keypoint size; whilst nadir orientated views contain the most significant geometric errors. Results suggest that reconstructions should be based on either the 'High Full' scenario or 'High Oblique' scenario, providing a trade-off between variance in the initial sparse cloud and consistent surface geometry. Using only oblique images appears to provide the best results, however this poses additional dangers associated with over obliquity and increasing the percentage of image beyond the FOV.

The difficulties of using only oblique imagery to provide catchment-wide reconstructions are effectively evidenced in a simple trial. For this, an alignment was created of the full 35 km of the Dart using 6288 oblique images. Under High quality settings, only a total of 805 cameras were aligned, generating an incomplete sparse cloud comprising just 590,000 points with significant reprojection error of 6.4×10^{11} pix and significant incoherence (gaps and separation of the cloud). Whilst oblique imagery therefore appears to provide high quality results over a small study area, using this in combination with nadir photography is required to map larger extents using classical corridor flight paths. Sparse cloud noise may be reduced under such a survey strategy through the implementation of a mask to eliminate peripheral regions prone to large pixel distortion and loss of sharpness. In the interest of increasing automation/standardization of the process this remains optional, partly due to the varying mask requirement depending on effective image angle and the redundancy of the sparse cloud when densified within APS.

6.2.2.5. ***Summary***

Results from the analyses described above support the use of convergent imagery and the densification algorithms within APS. Whilst increasing baseline distances may reduce sparse cloud noise, there is minimal impact once dense cloud construction has occurred. Moreover, maintaining the high image density builds-in redundancy into the survey which may prove important in areas that are poorly mapped such as the gorge.

6.2.3. **Image Processing: Implementing camera GPS and correcting for error**

Chapters 4 and 5 identified the potential benefits of using a camera GPS for large aerial surveys, which relate principally to constraining the image search during the SfM bundle

adjustment. For the dual camera strategy employed here, the use of camera GPS is not just advantageous, but essential, as this restriction on the bundle adjustment is required to manage the computational burden associated with doubling the image numbers. The 2015 survey used two GPS receivers, one connected to each camera, with the 3D position recorded in WGS84 coordinates written on the associated EXIF of each image. As described above, the Solmeta receivers used here are low cost, code-only systems and their use in the field setup posed some significant challenges. The use of low grade receivers is compounded here by the high speed of recording required (1Hz) and flying speed of 90 knots used, which implies that small errors in position or time-averaging are likely to magnify into significant changes in estimated camera pose.

Performing an initial reconstruction of the Dart in 2015 using the available camera GPS data resulted in the construction of a poor quality sparse cloud. The cloud itself was prone to numerous areas of incoherent cloud geometry (see Figure 6.9), whilst the entire block contained a directionally consistent mean camera error of > 160 m. This camera error as provided by APS is the difference between the *a priori* positional data and the posterior position once the ground control network has been incorporated into the bundle adjustment. This error essentially relates to the difference between where the GPS proposes the camera to be against where the bundle adjustment proposes the camera to be. Despite the use of a low grade receiver, errors in the order of 160 m are symptomatic of an inherent issue in the data collection process.

An initial analysis identified systematic translational errors of the reconstruction in the order of >160 m with regions failing to construct due to spatial inconsistencies in the sparse cloud. To eliminate the camera errors and enable accurate reconstruction, a heuristic workflow was employed to address the impacts of improper camera estimation via the on-board GPS.

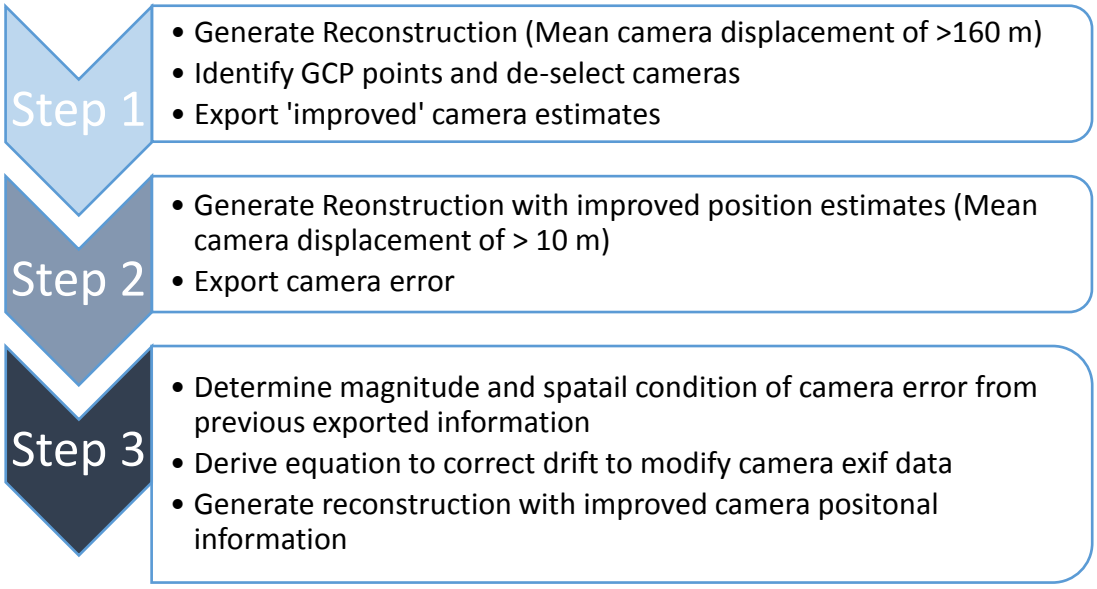

Figure 6.8 Heuristic process applied to address initial inaccuracies in camera GPS exif data derived from on-board GPS.

6.2.3.1. ***Step 1: Translational Error***

The initial reconstruction from the 2015 Dart survey was performed within APS following the approach detailed in section 3.4, under the High alignment setting alongside *a priori* camera positional information as provided by the GPS. The reconstructed sparse cloud was projected in NZGD 2000 from the *a priori* positional data before being refined through the introduction of a total of 87 GCPs distributed throughout the reach and subsequent optimization. The introduction of the ground control network provides the basis from which APS is able to estimate the camera error (posterior position – *a priori* estimate). As noted previously there was a significant degree of camera error and regions of incoherent geometry within the cloud; the consistent (translational) camera error shall be addressed here.

For practical reasons, namely time constraints, a series of steps to address the camera error were undertaken using the Low alignment setting within APS. The initial sparse cloud under the Low alignment settings generated a point cloud of ~ 9 million points, once more exhibiting significant areas of incoherent cloud geometry creating breaks in the point cloud. Once more the addition of GCP data revealed an average horizontal camera error in the order of >160 m consistent across all imagery. To address this, the camera network was de-selected within APS and the reconstruction optimized based

only on the exterior constraints provided by the ground control network. The resulting point cloud provides updated estimations of the cameras posterior position, which may be exported to provide the seeding exif data for subsequent reconstructions.

6.2.3.2. ***Step 2: Camera Drift***

Updated reconstructions using the improved camera positions derived from step 1 reduced posterior camera errors to a mean of <20 m, however there remained several persistent areas where the point cloud failed to align coherently as illustrated by the example shown in Figure 6.9.

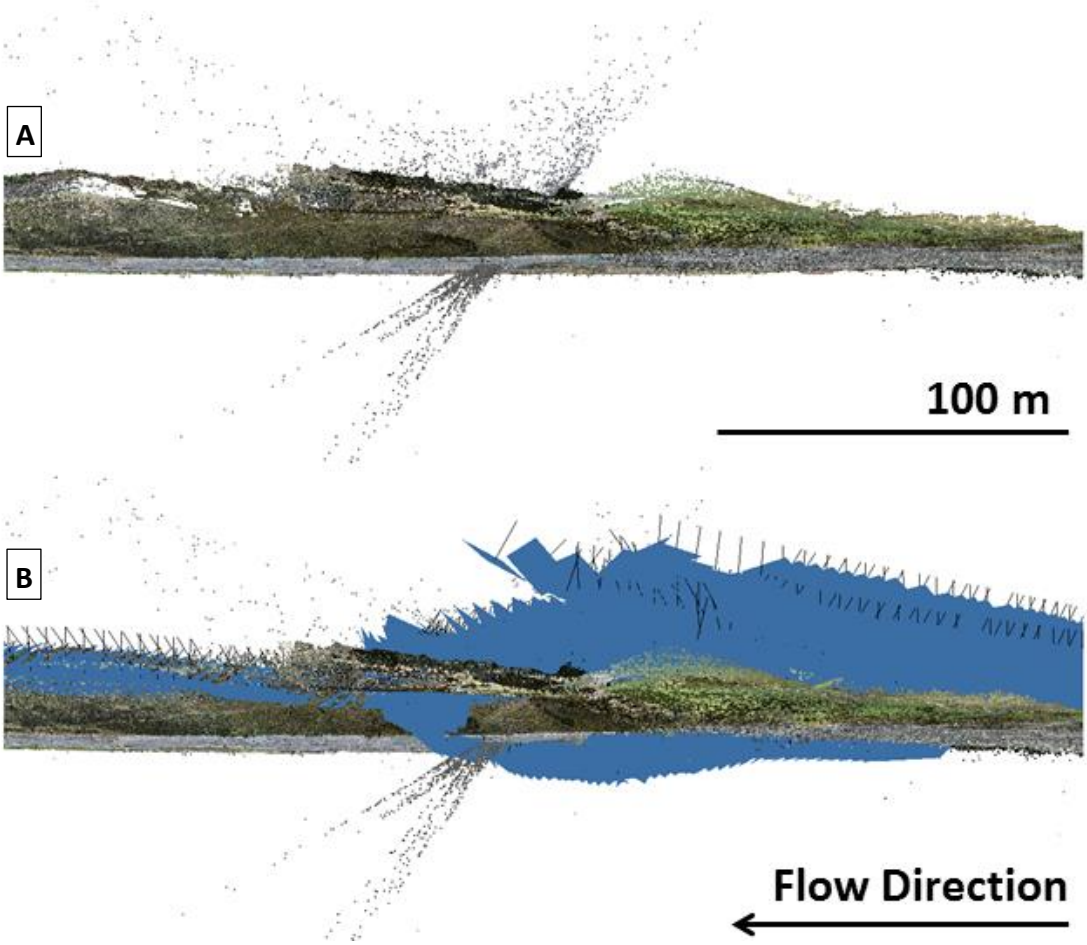

Figure 6.9 A) A region of incoherent cloud geometry occurring within the sparse cloud within proximity to paradise and the corresponding camera projections within APS. This incoherence takes the form of a planar intersection of the dominant cloud orientation. B) This divergence in the point plane can be seen to coincide with a divergence in the camera projections.

A number of these artefacts were significant and could not be rectified by manually realigning the cameras. On closer examination, it was clear that in these areas, the

estimated camera positions determined by APS exhibited significant and abrupt changes, in particular in terms of their Above Ground Height when heading from right to left. This raised the possibility that camera error was increasing systematically along lateral and longitudinal transects, resulting in divergent positional estimates which impacted on the subsequent reconstruction quality.

6.2.3.3. ***Step 3: Correcting Drift***

To determine the cause of these artefacts a subset of 65 images in an error prone region proximal to Chinaman's Bluff were used to assess the origin of increasing camera errors along a transect. The 65 images were collected for a region central to a flight path transect and as such it is presumed that the cameras in this analysis are already subject to prior accumulated error. From this point on, the camera error is assumed to be a result of GPS error, whereby *a priori* misestimation is a direct result of the GPS. By plotting the 3D error for the 65 cameras along both longitudinal and latitude direction we are able to determine if there is a systematic offset present in GPS error relating to flight lines or if these are random. Figure 6.10 provides

To determine the presence of a trend, the 3D GPS error (posterior camera position - *a priori* camera estimate) was plotted for each camera along both longitudinal and latitudinal directions of the flight transect, Figure 6.10.

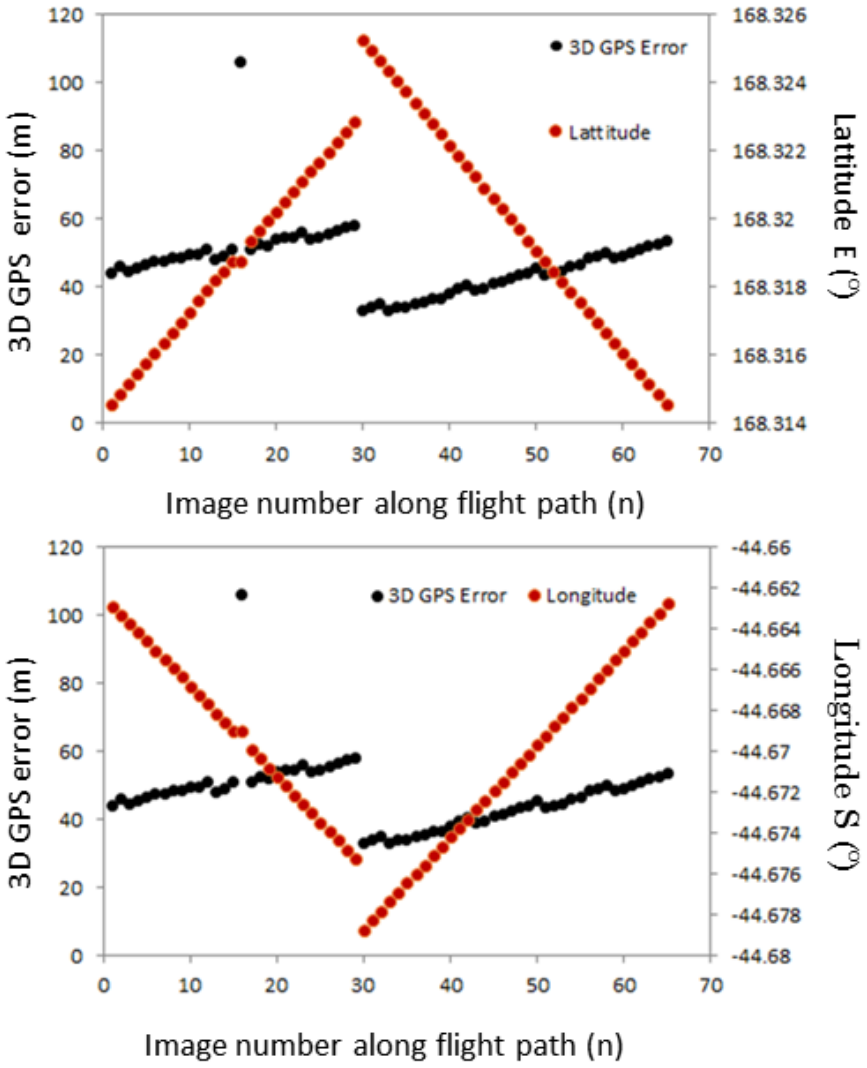

Figure 6.10 Camera 3D GPS error and relative X and Y position plotted against photo number. The photo number corresponds to the order captured whilst travelling along a transect.

From this analysis, the *a priori* GPS error appears to be linearly related directly to the flight path direction, generating a near constant rate of increase in error with distance. The abrupt break in the GPS error occurs where the flight path extends beyond the region of interest covered by the 65 images. As such, this represents the point at which the capture of imagery continues from a different flight transect. The change in error with distance along transect was plotted for both latitude and longitude to determine the X and Y component of the lateral camera drift per unit metre. Given the linearity of the drift, a simple regression analysis can be used to correct the estimated camera position which leads to the following equations:

$$C_{lat} = 0.21x - 15 \quad [6.1]$$

and

$$C_{long} = 0.55x + 28 \quad [6.2]$$

where, C_{lat} and C_{long} are the camera corrections applied for both latitude and longitude, x is the number of images along the transect and integers reflect the prior accumulated error along the flight path. Equations [6.1] and [6.2] provide a means to correct the X and Y vectors of drift in the GPS error for the flight path directions presented for the test region of 65 images and by extension the ~ 5 km flight section they are part of in part of the wider reach reconstruction. Therefore each equation must be tailored based on the orientation of flight transect for application over the entire survey area.

*6.2.3.4. **Application of the camera correction***

Using the method described above to re-estimate the a priori (GPS) camera position resulted in a reduction of the posterior estimated error from 80 m to 7.9 m. Furthermore the point cloud was continuous without artefacts. This process was therefore applied to the whole camera set using tailored versions of the drift equations described above specific to differing regions of flight lines and accounting for turning locations. In the turns at the end of each flight line, the reduction in helicopter speed is used to reset the correction equations (6.1 and 6.2) for the next flight line, starting with an accumulated error and distance of zero. In contrast, the values derived from Figure 6.10 are mid-flight transect and so incorporate an integer value for previously accumulated error.

Initial tests of the method across the entire study area were found to reduce the average camera error to 12 m with no regions of incoherent cloud geometry. Subsequently, camera estimates were improved further by the re-application of step 2, whereby the bundle adjustment was optimized via only the exterior constraints provided by ground control. This provided an updated set of co-ordinates for incorporation into the final bundle adjustment of the 2015 survey, which resulted in a posterior position – *a priori* estimate error value of just 2.5 m.

The large initial positional errors were attributed to GPS latency associated with low cost GPS units. There is limited documentation associated with the processing of position, but it is possible that these errors are most likely associated with a degree of temporal averaging of positions, or that estimates are partially auto correlated to smooth plotted trajectories. While the approach probably works well for low velocity (ground based) motion, it is likely to become problematic at the speeds used in this survey. The effects are potentially compounded by increases in velocity along each flight line as the helicopter accelerated at the beginning of each turn (Bowen and Danya, 2014).

While the introduction of the camera GPS failed to provide the initial precise and accurate positional data desired to inform the exterior camera pose estimation, the process to engineer accurate positions described above does require fewer inferences than the introduction of estimated GCPs used in 2014.

6.2.4. **Final reconstruction of the 2015 survey**

Following the implementation of the GPS workflow a final reconstruction for 2015 was performed using an upgraded Dell workstation with dual Xeon 3.4 MHz processors, each with 8 cores; 512 GB RAM; and a 2 GB Graphics card. The updated camera co-ordinates were used to inform a reconstruction performed under the High alignment and reference settings within APS, specifying a 40,000 keypoint limit and a limitless tiepoint limit. Processing time to complete an alignment within APS was 11.5 hours, producing a sparse cloud of 4.5 x 10^7 tiepoints over an area of 80.4 km^2 with a reprojection error of 0.19 m (0.86 pix) and mean keypoint size of 3.83 pix. The sparse point cloud was refined using the gradual selection tool to remove points associated with high projection errors (RPE) with an initial threshold of 1 pixel selected before iteratively being reduced until 0.8 pixels. The RPE is associated with incorrect feature matches (see section 3.4.4) and is advisable to be reduced before performing the optimization step within AS. This thinning of the cloud resulted in a final sparse cloud of 3.2 x 10^7 points. The dense cloud consists of ~1.3 x 10^9 points with an average point density of 17.5 per m^2 which is deemed sufficient to generate DEMs of 1 m resolution, see Figure 6.11.

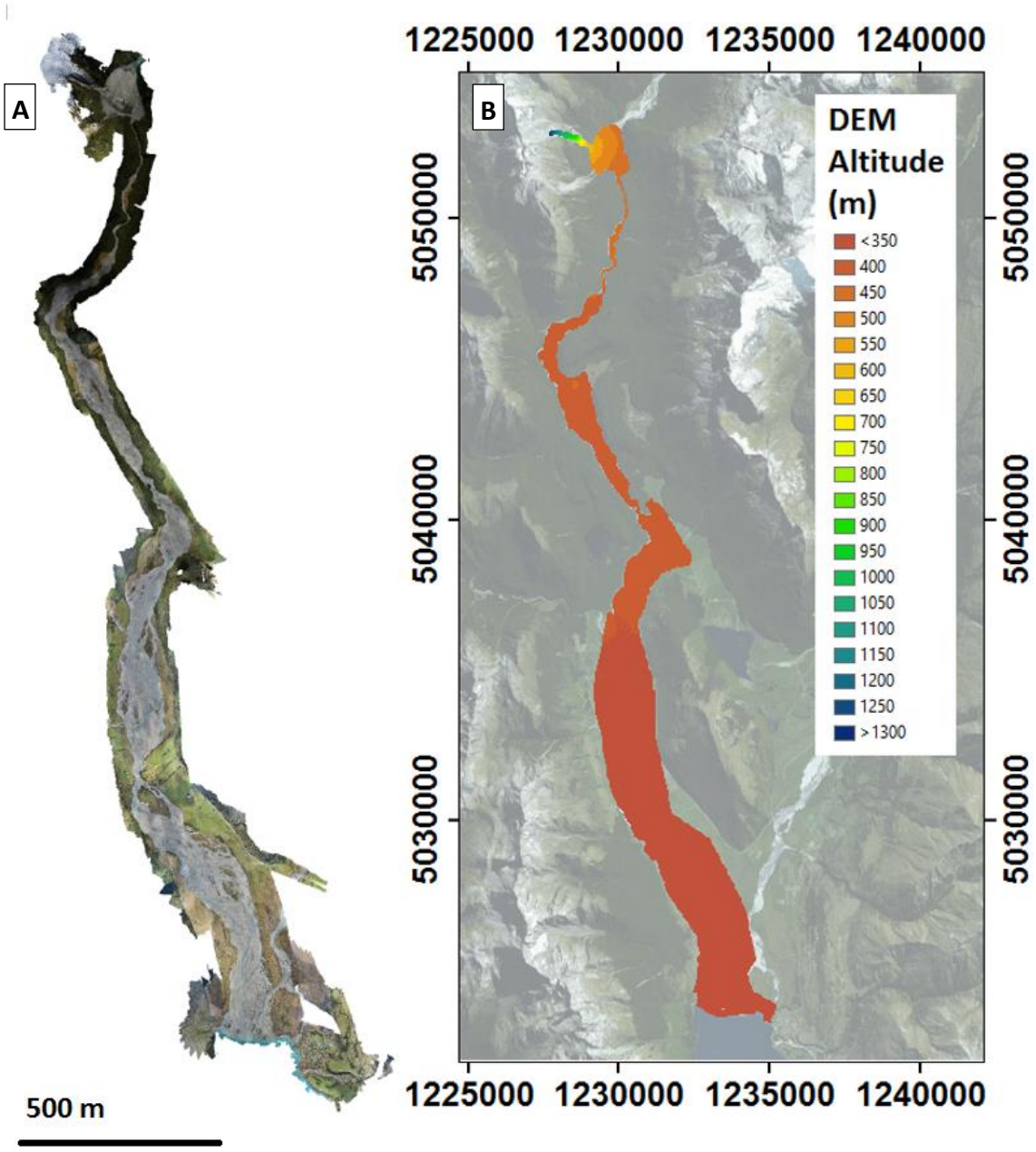

Figure 6.11 A) Dense reconstruction of the 2015 catchment consisting of 1.4 x 10^9 points and B) the DEM produced within APS. *Projection: NZTM 2000*

In line with the strategy adopted for the 2014 survey, the 2015 survey was geo-referenced using a network of manually placed GCPs comprising of the same 1 m^2 red tarpaulin squares, however this time marked with a central 10 cm black square to aid identification of the mid-point within APS. The GCPs were deployed and surveyed over a period of four days, prior to the aerial survey with 3D coordinates obtained in NZTM using a combination of RTK and PPK GNSS observations. As previously, these

observations were tied to a local network of 1st order reference stations which in turn were localized from the Land Information New Zealand (LINZ) Position NZ network of GNSS Continuously Operating Recording Stations (CORS); see Chapter 3 for further details.

The GCP distribution represented the best attempt at a quasi-uniform distribution throughout the catchment, however, issues associated with the landscape prevailed. Once more control was limited to a) stable regions upon the floodplain unaffected by episodes of rising stage; b) publicly accessible land and; c) regions with a clear line of sight to the helicopter, particularly relevant in the canopied areas of the gorge. Due to an episode of peak flow representing a flood event of a 2 year return period magnitude a number of targets required re-positioning and surveying. This resulted in a number of GCPs located at a greater distance from the channel banks during the 2015 survey in relation to 2014. Table 6.5 provides the residuals for all 87 GCPs in the 2015 survey.

Table 6.5 GCP residuals for the 2015 Dart survey and 2014 for comparison, where 2015 GCP (n = 87) and 2014 GCP (n = 80).

Statistic	Total (m)	X (m)	Y (m)	Z (m)
2015 Mean	0.05	0.00	-0.00	0.00
2014 Mean	0.06	0.01	0.00	0.00
2015 MAE	0.05	0.02	0.02	0.02
2014 MAE	0.06	0.03	0.03	0.04
2015 SDE	0.04	0.030	0.03	0.05
2014 SDE	0.10	0.06	0.034	0.09
2015 MaxE	0.32	0.08	0.14	0.28
2014 MaxE	0.62	0.39	0.17	0.58

When undertaking an analysis based upon GCPs, the caveat remains that it is important to recognize that residual errors incorporate a bias due to their role within multiple stages of the SfM workflows, i.e., as exterior constraints in the bundle adjustment; targets used to define a posterior registration of the sparse cloud; or as targets to optimize the exterior and interior orientation. Furthermore, the GCPs themselves are not without uncertainty and inevitably incorporate errors associated with the raw GNSS.

The improvement between surveys is well-represented across all metrics for GCPs in Table 6.5. The greatest GCP residual differences occur in terms of the SDE and ME, and reflect both the improved precision and accuracy of the 2015 survey. The MaxE value

for the 2014 survey is 0.62 m within the 2014 survey, and this large anomaly actually reflects a GCP used during optimization, implying the presence of severe systematic errors that could not be corrected with exterior constraints. By comparison a MaxE value of 0.32 m in 2015 reflects a more accurate surface which is less impacted by large variations from reality within the bundle adjustment. The distribution of GCP residuals may be seen in Figure 6.12.

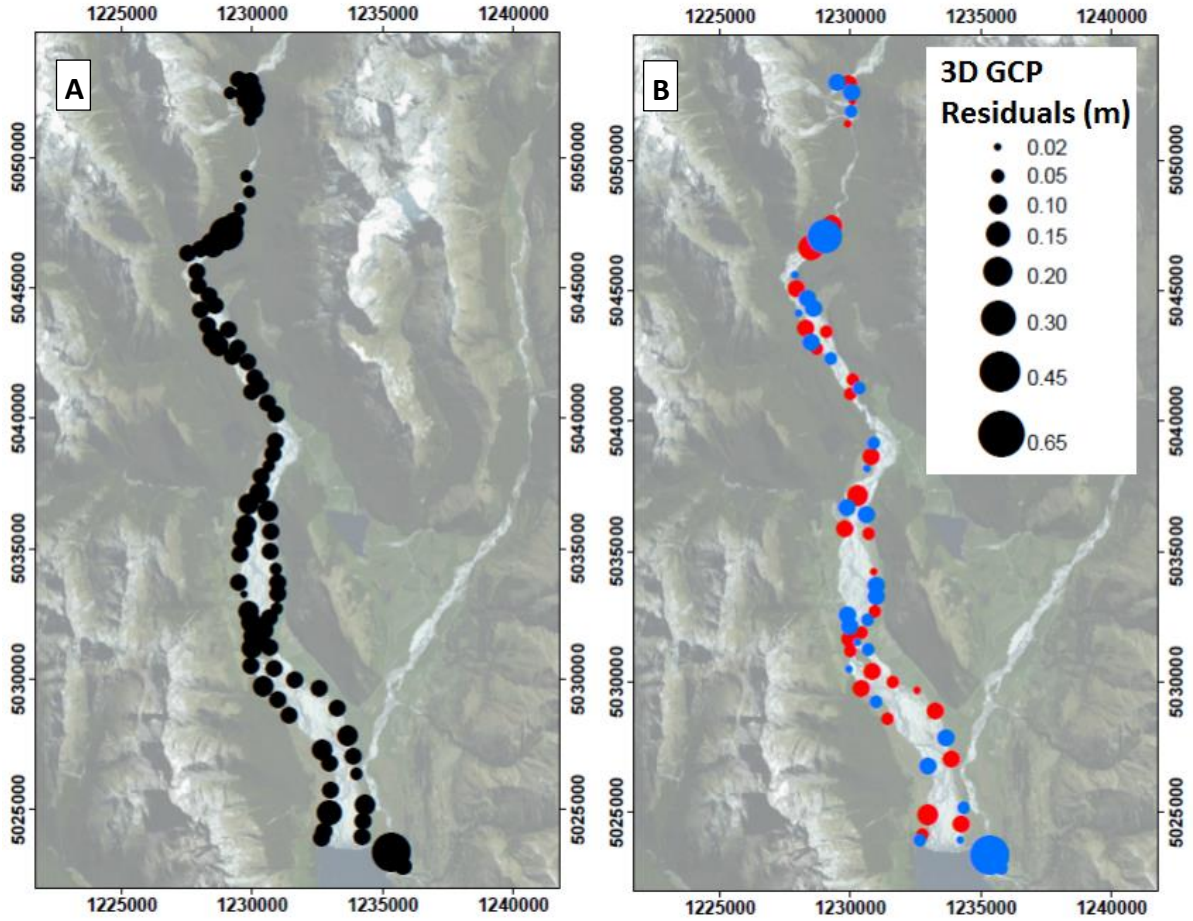

Figure 6.12. A) 3D GCP residuals for the 2015 survey and B) Positive (blue) and Negative (red) residuals for the 2015 survey. *Projection: NZTM 2000*

It is clear from this, that the 3D GCP errors are less than 0.2 m across most areas of the model, with notable exceptions in both Glenorchy and within the Gorge. These are reflected in the distribution of vertical errors, with a positive errors recorded in both locations. The scale of errors in comparison to the 2014 survey is significantly reduced, with a lack of errors observed at the Hillocks and Paradise in 2015. The GCP analysis here

provides a description of the 2015 survey however independent control is required to fully assess the model accuracy.

6.3. The 2015 Model evaluation

6.3.1. ICP residuals

In order to provide a first assessment of model quality a simple leave-p-out cross validation test was performed. This involved sub setting the GCPs, extracting 20 of the 87 GCPs, distributed evenly across the study site to provide independent check points (ICPs). The remaining 67 targets were then used to register and optimize the model once again (see section 3.7). The 20 ICPs where separated by a linear distance of typically 2-3 km (see section 3.5 and Figure 6.12). Their location was chosen to reflect, as closely as possible, the network of ICPs derived for the 2014 survey with 19 of the 20 within 500 m of the 2014 ICPs. A more robust, leave-one-out cross-validation was not possible due to the runtimes involved creating 87 individual alignments using the limited computer facilities available.

Residual statistics for horizontal, vertical and total positional errors were calculated for reconstructions based on the network of 67 GCPs and 20 ICPs. The ICP values provide independent assessments that are consistent in location and number to those used in 2014. The results for both 2015 and 2014 are presented for comparison in Table 6.6.

Table 6.6 The relevant ICP residual statistics for both the 2014 and 2015 Dart catchment reconstructions.

Configuration	Statistic	Total (m)	X (m)	Y (m)	Z (m)
2014 ICP	ME	0.44	0.08	-0.12	-0.11
	MAE	0.44	0.15	0.18	0.33
	SDE	0.62	0.23	0.40	0.58
	MaxE	2.67	0.72	1.68	2.00
2015 ICP	ME	0.23	0.02	-0.03	-0.03
	MAE	0.23	0.08	0.07	0.18
	SDE	0.16	0.11	0.08	0.24
	MaxE	0.63	0.27	0.19	0.62

The first key observation, is that the Mean Error (ME) (eq. 3.1) in 2015, which captures systematic bias in the model is close to zero in each of the horizontal and vertical components (0.02, -0.03 and -0.03 m, x,y,z respectfully). While the MAE (average 3D displacement) sums to 0.23 m, the lack of bias in any specific dimension is encouraging and suggests a robust overall alignment of the model to the underlying GNSS observations. On closer examination of the ICP errors, it appears that vertical

dimension is responsible for the majority of the total 3D error, with significantly greater proportional contributions to the 3D MAE, SDE and ME.

The quality of results may usefully be benchmarked against the maximum achievable accuracies that could be achieved as discussed in Chapter 4. Empirical estimates of the likely accuracy based flying height ratios imply idealized accuracies in the range of 0.38 m (James and Robson, 2014) to 0.59 m (Smith and Vericat, 2015). By contrast, the theoretical approach of Luhmann (2003) suggests 0.52 m in response to the reduction in baseline distances in comparison to the 2014 survey. The observed accuracy of 0.23 m, measured in terms of the ICPs, therefore reflects a surprisingly good result. This should not necessarily be a surprise however, as Luhmann's (2003) analysis is based on stereo-pairs, as opposed to a pairwise multi-view approach used in APS. The maximum error recorded is marginally greater than that proposed within the literature and serves to highlight potential distortion of the model in areas with sparse ground control.

The 3D MAE is reduced by approximately 50%, from 0.44 m (2014) to 0.23 m (2015); principally reflecting a reduction in the vertical error of 0.33 m to 0.18 m. In addition to improvements in the MAE, the SDE, reduces by over 75% between the surveys, from 0.62 m to 0.16 m. Importantly, this suggests an improvement in the consistency of the solution across the area of interest, with fewer outliers indicating areas where the solution is significantly in error. This latter point, is also reflected by the lower MaxE recorded. The largest residual in 2014 measured 2.67 m, but is now reduced to just 0.63 m in 2015. These extreme errors are symptomatic of distortion of the model in the peripheral regions of the area of interest where there is sparse ground control.

The improvements in the quality, both in terms of the overall accuracy (ME, MAE) but also consistency (lower SD, Max E) of the model reflect the use of convergent imagery and the introduction of *a priori* camera positional information. However, it should be noted that the 2015 survey also benefitted from an improved distribution of ground control, enhanced image capture within the gorge, and fewer regions of high contrast. As with many natural experiments, these confounding influences make it difficult to fully ascribe cause and effect, but it is clear that in sum total, the modifications proposed following the laboratory studies in Chapter 5 have generated a more reliable model.

The statistical metrics of model quality described above provide only an overview of the model fit and fail to provide an insight into the spatial characteristics of these errors. As observed in Chapter 4, the 2014 reconstruction exhibited significant systematic errors that were characterized by strong spatial autocorrelation. The use of a convergent camera network was designed specifically to mitigate such errors by improving the estimation of both radial distortion and focal length. An insight into the spatial structure of errors can be gleamed by plotting the GCP (n = 87) and ICP (n=20) residuals spatially, as shown in Figure 6.13. Here, the model residuals as proportional symbols, and shown for both the magnitude of 3D error and the total and direction of the vertical (elevation) error, using different colours (red and blue) to indicate the direction of error.

The 3D ICP errors once more indicate the difficulties of reconstructing the Gorge along with sections of the delta. The Gorge is once again prone to larger errors despite improved image capture (vs. 2014). The region is difficult to reconstruct due to both the reduced active width, providing less observable bars with a greater percentage under water and by virtue of the surrounding dense vegetation. GCP placement was improved, however remains difficult in the region due to access restrictions with the necessary GPS equipment to obtain readings, or indeed be visible with the overhanging vegetation.

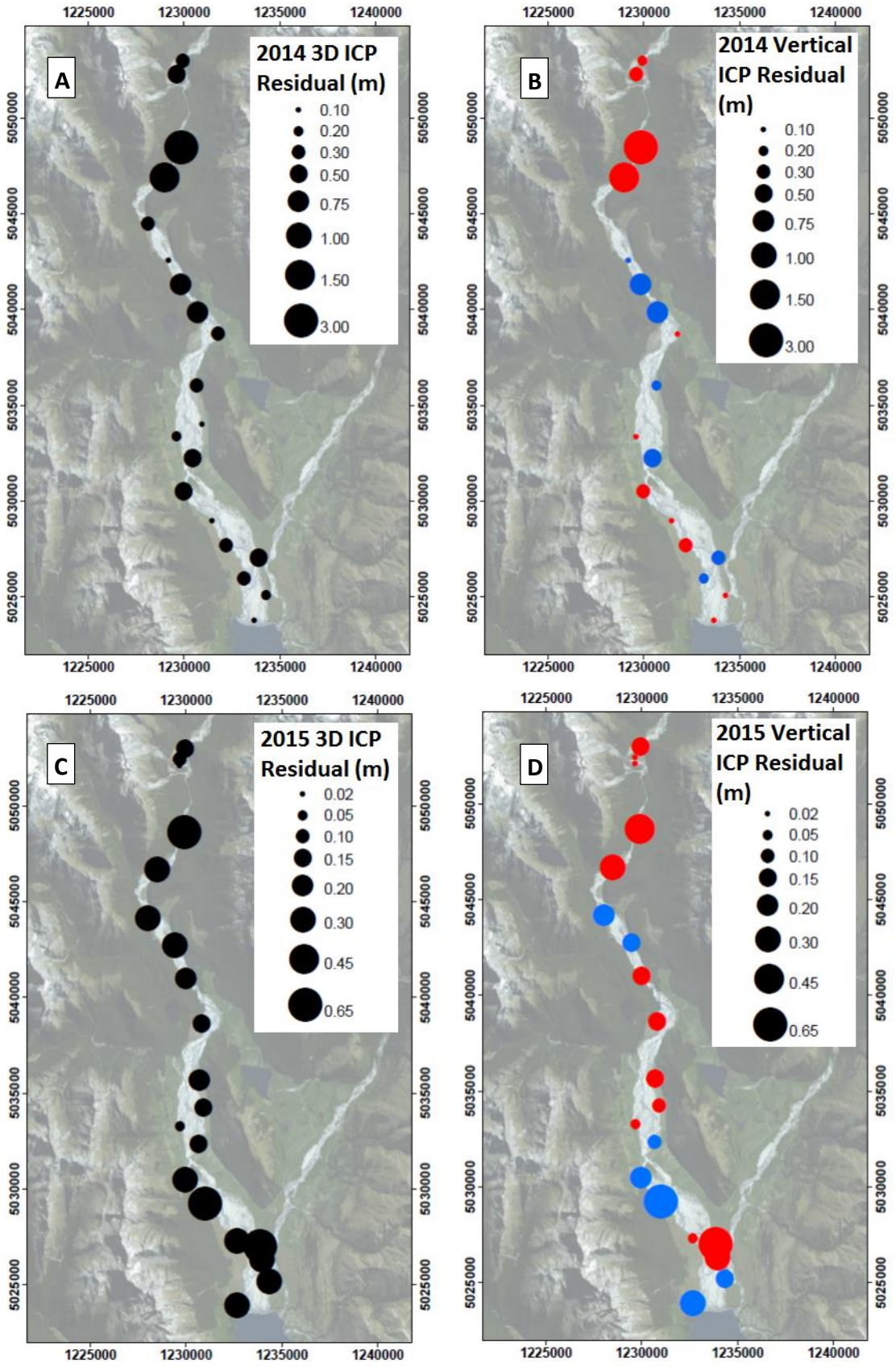

Figure 6.13 GCP and ICP errors across the Dart catchment for the 2015 survey, for both 3D and vertical error for the final High Alignment reconstruction. A) 3D GCP error, B) 3D ICP error, C) Altitude GCP error, D) Altitude ICP error. *Projection: NZTM 2000*

The largest ICP error is observed in the flat paddocks (grazing areas) on the true right of the braidplain near the Hillocks. Given the lack of supporting behaviour shown by the biased GCPs in this area, it is possible that this result is an anomaly and could reflect disturbance of the target between the GNSS survey and imagery capture or indeed a GNSS measurement blunder. Perhaps of more concern, is by plotting the direction of vertical change, there is some evidence of an oscillating pattern of positive and negative values, fluctuating longitudinally at a low spatial frequency (5+ km). Given the low density of ICPs it is hard to confirm such evidence, but this is somewhat reminiscent of the patterns observed in 2014 survey, albeit at a much reduced magnitude.

6.3.2. Detailed Model Assessment at Glenorchy

The difficulty of drawing reliable inferences about model quality given with the low density of ICPs is an inevitable but unfortunate result of the scale of analysis. One approach to relax this constraint is to densify the network of check point's post-hoc by measuring observable 'natural' targets. Unfortunately, there are few stable, easily identifiable features across the study area, however, one area where there are abundant marks that can be easily measured is in the township of Glenorchy. To support this analysis therefore, a dense network of additional ground marks was surveyed following the aerial survey, during a return visit to the site in August 2015 (see Figure 6.14).

Figure 6.14 Collection of RTK measurement of a road marking within the township of Glenorchy to provide extra independent control points.

A total of 28 marks were documented, principally road markings that have negligible vertical relief and could be easily identified from in the raw imagery. Coordinates were obtained in NZTM using RTK GNSS observations.

This dense network of independent check points provide the opportunity for a highly focused assessment of model quality, albeit only over a small area of the whole model and one that is in the periphery of the broader area of interest. The results obtained for these new check points are show in Table 6.7.

Table 6.7 Comparison of ICP errors from the 2015 catchment wide reconstruction compared to a localized specific Glenorchy reconstruction.

Reconstruction	Mean Vertical	MAE 3D (m)	3D SDE (m)	Max 3D (m)
Catchment Wide	-0.03	0.27	0.12	0.46

As with the analysis of the wider distribution of check points, there appears to be no significant bias, with a low mean vertical error of -0.03 m. Moreover, the 3D MAE of 0.268 m is also consistent with the wide area measurement of 0.27 m and the SDE is actually lower at 0.12 m compared with 0.16 m (perhaps reflecting the solid, low relief surrounding surface). While these results need to be treated cautiously given their localized nature, it is particularly encouraging to see such strong model perform close to the margins of the photogrammetric block.

6.3.3. The effect of optimization of the model structure and systematic errors

A key aim of the 2015 survey strategy was to limit reliance upon post-hoc optimization of the bundle adjustment using ground control as exterior constraints. Whilst this process can lead to significant improvements in the estimation of the intrinsic camera parameters and pose, it also creates the potential to introduce systematic spatial errors that are difficult to mitigate without the addition of more and more ground observations. The sensitivity of the model to the corrections introduced during optimization can be determined by comparing the pre- and post-optimization models using a DoD approach as used in Chapters 4 and 5. This was achieved by creating DEMs of the point cloud, following the same process used on the 2014 datasets. First, the point cloud is decimated to provide a min elevation over a 5 x 5 m window using the Topographic Point Cloud Analysis Toolkit or ToPCAT (Brasington et al., 2012). This

reduced resolution sparse cloud is then interpolated using a Delaunay Triangulation to derive a Triangular Irregular Network (TIN), which in turn, is then linearly resampled to a raster DEM at 5 m resolution. This process was applied to both pre- and post-optimization point clouds to generate comparable DEMs that could then be subtracted to yield a DoD. The result is presented in Figure 6.15, which shows both the correction inferred in 2015 and 2014 for comparison.

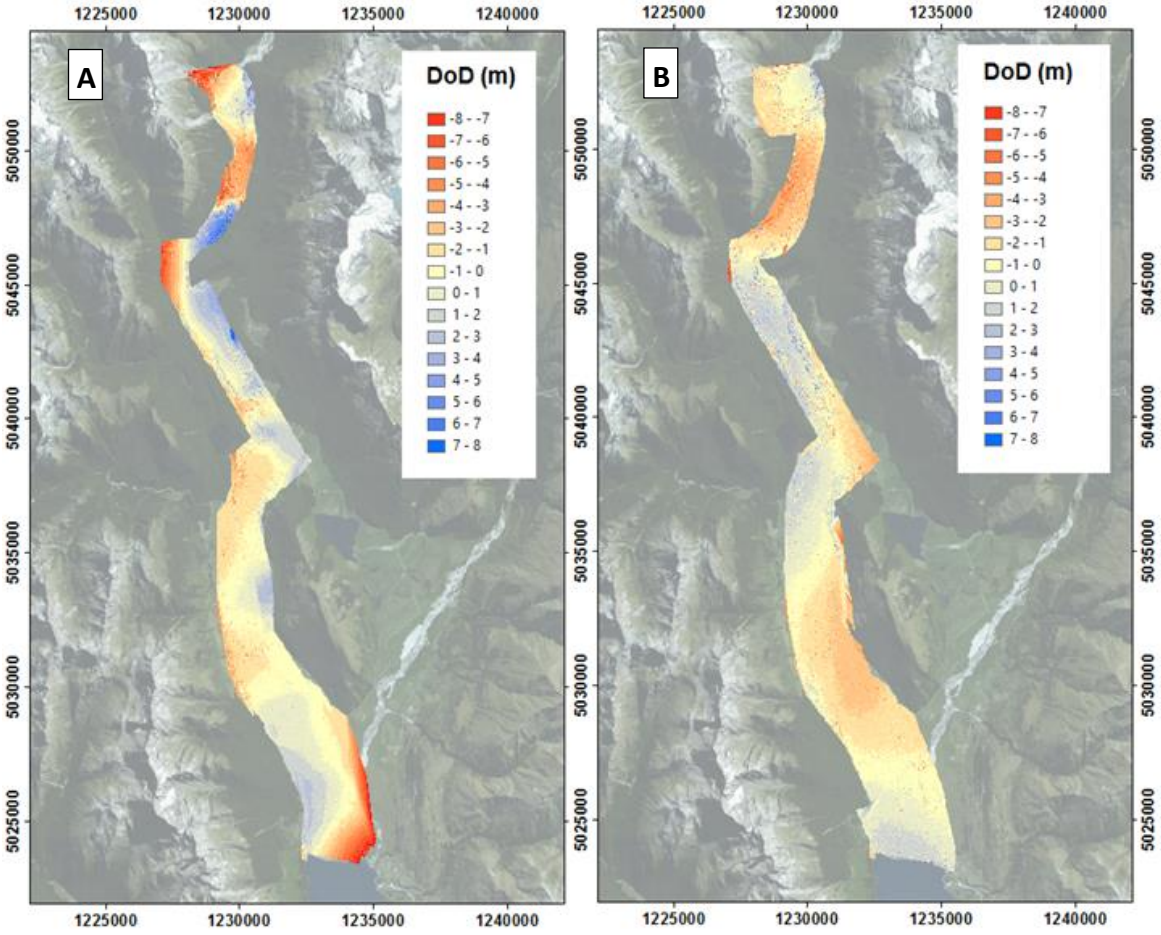

Figure 6.15 Optimization DoD (5 m DEM resolution) for the 2014 and 2015 catchment reconstructions within APS. The DoD is generated by subtracting the initial sparse cloud DEM from the optimized DEM to provide an insight into the structural change occurring during this process in APS. *Projection: NZTM 2000*

The spatial correction applied during the optimization phase is greatly reduced for the 2015 survey, with >68% of the model changing by less than 2 m, by comparison to only 50% in 2014. Most significantly however, the extent and severity of correction appears to be significantly lower overall and there are now areas with change of +/- 6 m by contrast to 2014. The elimination of these spurious regions, along with the general reduction in systematic change signifies an improvement in both accuracy and precision

in 2015 achieved through the adoption of revised survey strategies. Despite the overall improvement, the correction pattern does once again exhibit strong spatial autocorrelation, as hinted at in the ICP analysis. This is manifest as an oscillating, or wave pattern, with positive to negative corrections alternating longitudinally, mirroring the corridor flight paths. The amplitude of correction is significantly lower however, of the order of 1-3 m and the wavelength of correction pattern appears to have increased (from ~ 5 km to ~ 10 km). The Gorge reach once again exhibits the largest degree of change, ranging between 2- 6 m and likely reflects the sparse GCP coverage and the low reliability of keypoints in this largely inundated and forested region.

As described in Chapter 5, the degree and direction of optimization correction was found to be closely correlated to magnitude of resulting camera displacement. The respective displacement for 2014 and 2015 is shown in Figure 6.16. In this analysis, it is clear that camera displacements following optimization is significantly lower in 2015, typically of the order of < 1 m, whereas displacements of > 5 m were typical across the whole model for the 2014 survey.

It was hypothesised in Chapter 5 that the wave like nature of the optimization correction is a result of the interaction between the solution of the interior and exterior camera geometries. Due to the challenges faced during the initial construction of the 2014 survey, systematic errors in camera location - a result of an initial seeding of the bundle adjustment with low quality information – appeared to dominate the resulting pattern. The result was seen in significant compensation between the K1 and focal length parameters, and the exterior pose (in particular affecting height). The absence of such major interactions evident in this revised survey, point to the influence of the camera GPS and convergent image network providing a more robust and accurate initial estimation of the optimal structure from the initial SfM bundle adjustment (without GCPs as constraints).

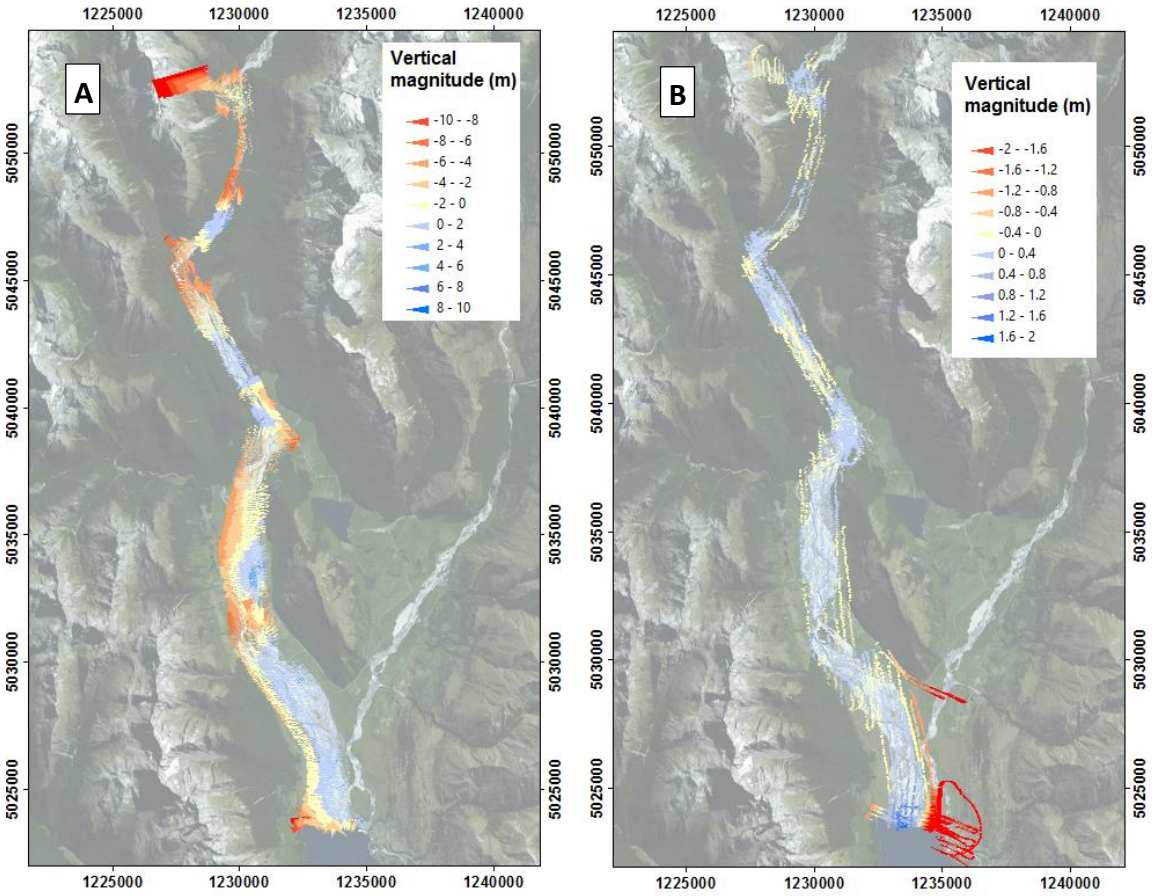

Figure 6.16 Camera displacement during optimization for 2014 and 2015. The size of arrow is a scaled representation of x/y translation and the colour refers to the magnitude of vertical change. Note the difference in scale between years. *Projection: NZTM 2000*

This interpretation is corroborated by analysing the change in the key intrinsic and exterior parameters of the model as shown in Table 6.8. This shows the percentage change in the parameter of the structure measured before-after optimization and is tabulated for both the 2014 and 2015 surveys respectively.

Table 6.8 Values presented for changes to both internal and external calibration parameters during optimization change within APS.

Year	Focal length change (%)	K1 change (%)	Initial GCP error (m)	Initial camera error (m)	Mean camera elevation change (m)
2014	1.79×10^{-2}	1.58×10^{-2}	2.92	3.88	1.03
2015	1.04×10^{-2}	0.15	1.66	2.48	0.11

Theoretically, the statistics presented in Table 6.7 should represent reduced values of change during the optimization phase for the 2015 reconstruction. This holds true for all

values aside from the radial distortion parameter which appears to be an aberration as far as 2014 is concerned. This is likely a result of the identification of an incorrect local minima during the 2014 calibration phase in response to the degree of error present in the initial cloud. The value for 2015 of 0.15% is also lower than expected, however broadly in line with previous estimates of change within a convergent network (0.28% for Glenorchy). The reduced degree of change may be a resultant factor of both the corridor geometry and the greatly increased GCP numbers in comparison to previous test scenarios. The degree of change required to focal length for 2015 is decreased which is reflected in a reduction in the mean elevation change of cameras during optimization, whilst the surface of the 2015 reconstruction is also closer to the GCP layout.

6.3.4. Benchmark Comparison: UAS survey of the Hillocks

The analysis of GCP/ICP residuals provides a point based sampling of the quality of the resulting SfM reconstruction and precludes a comprehensive, spatially explicit analysis of model quality. Benchmarking the model against an independent spatial model of the site, therefore provides a useful alternative assessment, even though such models inevitably incorporate their own uncertainties. Previously, this approach was used to test the model of the debris fan at Slipstream in Chapter 4 using a DEM derived from long-range TLS acquired as part of this project in 2014. Given the frequent disturbance of the fan between 2014 and 2015, these data provide an unreliable source of information to evaluate this more recent survey.

As such, an independent UAS-SfM of a 1.5 x 2 km site at the Hillocks was commissioned for this research and undertaken by an experienced photogrammetric team led by Pascal Sirguey from the National School of Surveying at the University of Otago. This model was based on a total of 2340 overlapping images (73% side lap and 80% end lap) covering a 2.8 km^2 area captured using a Trimble UX5 Remotely Piloted Aircraft System. The model was geo-referenced using 27 GCPs, surveyed using RTK GNSS tied to the LINZ national network of CORS (see description in Chapter 3 for further details). Photographs were acquired in two flights, comprising 39 lines in a cross flight pattern, orientated down and across valley. Images were obtained at an AGH of 122 m using a Sony NEX-5R 16 MP (15 mm, f/4.5). The imagery was processed in Trimble Business Centre (TBC)

v3.40, to generate a point cloud that was filtered to give ground points and interpolated into raster DEMs using natural neighbours, at a resolution of 0.15 m. A network of 15 check points were used to provide a quality assessment, revealing a 3D RMSE of 0.09 m and similar horizontal and vertical components of error (0.045 m, 0.077 m and 0.030 m for ENZ respectively). This should be treated as a conservative estimate of error as doubts surrounded the potential disturbance of two checkpoints. With these removed from the analysis, the 3D RMSE falls to 0.066 m.

As a comparison, the dense cloud for the whole study reach reported here had an average point spacing of 0.20 m. To facilitate the ease of DEM construction both raw point clouds, were therefore resampled to give orthogonal (aligned) 0.5 m raster resolution DEMs suitable for direct comparison using a DoD approach. The resulting DOD is shown in Figure 6.17, implemented by subtracting the model developed using the helicopter SfM from this new UAS survey (i.e., UAS – helicopter). Statistics reporting the match between surfaces are displayed in the inset table.

The comparison shows that over 63% of the coincident area lies within $\pm$0.25 m and 40% within $\pm$0.1 m. The MAE is 0.52 m, however this is inflated due to the presence of errors relating to vegetation (trees) which are not well matched, due to differences in the look angles, flying height and ground resolution of the photography. Additionally, large areas of the model are covered by pasture, representing a rough, grass surface. The wetted channel also fails to reconstruct consistently between surveys as expected due to the inherent difficulties of accounting for channels of depth or containing suspended sediment. The ME value of 0.11 m may initially seem indicative of a bias between surveys, however, once the channel network is masked the value drops to 0.04 m.

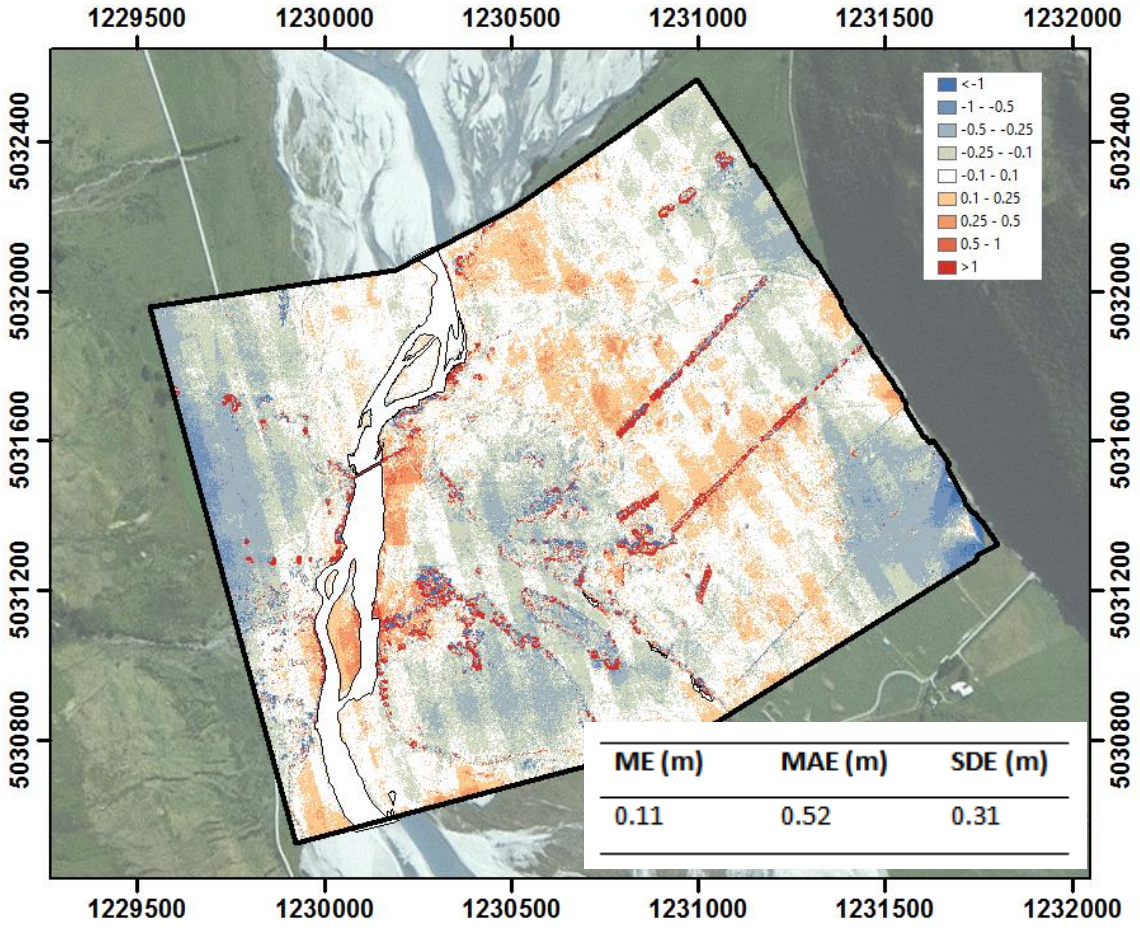

Figure 6.17 Dem-of-Difference of UAS-SfM survey minus catchment SfM survey located over a 2.8 km^2 area of the hillocks region of the Dart catchment. The main channel network is masked, to delineate the pattern of change observed. *Projection: NZTM 2000*

There is a systematic pattern of deformation between the two surveys that is observed as a ring of positive bias orientated centrally. The source of this is unclear as it correlates to neither ground control placement across surveys, however the relative symmetry within the survey region would indicate that this is a product of the UAS-SfM survey. The helicopter based survey would, by contrast, be expected to exhibit a systematic pattern that is not confined to the region. For corroborating evidence for this bias stemming from the UAS survey is that the regions exhibiting greatest difference are located peripherally beyond the boundary of UAS-SfM control.

The DoD in Figure 6.17 highlights an additional systematic pattern between surveys that mimics the flight path and photo collection of the UAS survey. This results in a pinstripe pattern of relative positive-negative change dominated by the vertical flight lines in the order of 0.15 m. The source of this error is determined to be an issue relating to the

aero-triangulation procedure in Trimble Business Centre 3.40 (TBC) processing software, which has been rectified by updated algorithms introduced in v. 3.61 (P. Siguey, pers. comm.).

These striped artefacts are clearly visual in a hillshade rendering of the surface as shown in Figure 6.18.

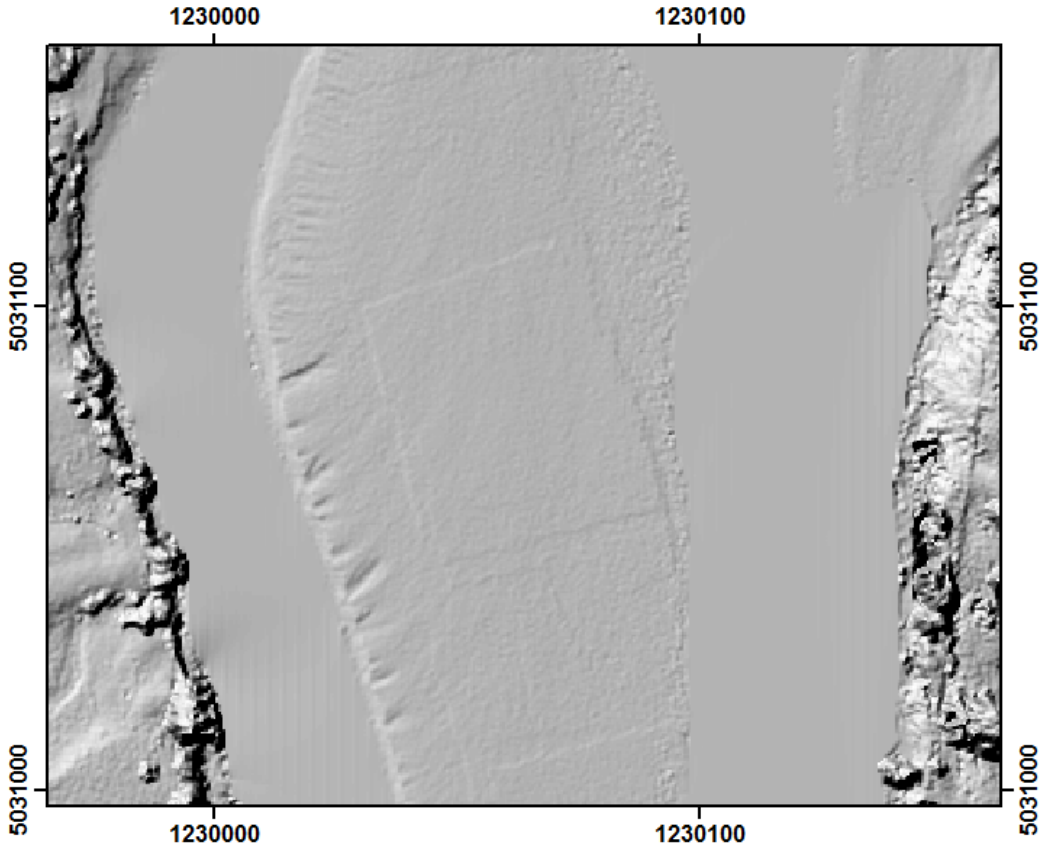

Figure 6.18 Hill shade of the DEM from the UAS-SfM survey, highlighting the presence of surface artefacts in the centre of the bar that do not relate to any natural geomorphological features. *Projection: NZTM 2000*

This bar provides a relatively flat or gently sloping surface for comparison, however there is a clear regular pattern superimposed on the DEM which, reflects the image footprint of UAS-SfM survey. It is difficult to determine whether this pattern is reflected across more topographically complex surfaces as vegetation and relief dominate. A similar pattern associated with changes in the overlapping footprint of images is also

evident, albeit to a lesser magnitude for the surface of the catchment wide DEM, see Figure 6.19.

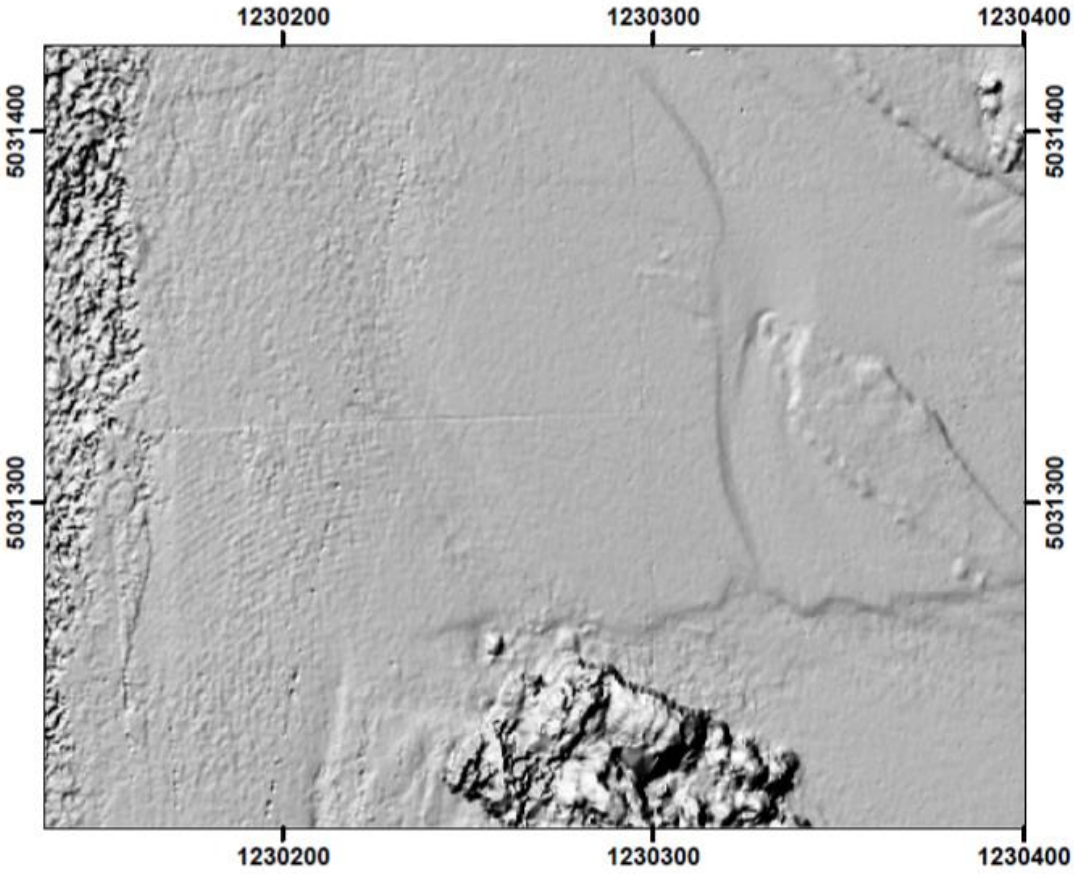

Figure 6.19 Image outline within the Hillshade for the catchment wide DEM. The detail is a lot less pronounced than the UAS-SfM upon the DoD. *Projection: NZTM 2000*

Such patterns clearly illustrate the difficulty of comparing surveys associated with different camera networks when subject to such overlap artefacts. This issue will persist regardless of the platform or strategy used and would require more complex surface adjustment to manage these problems and is beyond the scope of this research.

A further set of artefacts are visible through the inspection of the hillshade models, and as shown in Figure 6.20 below, appear to relate to shadow effects.

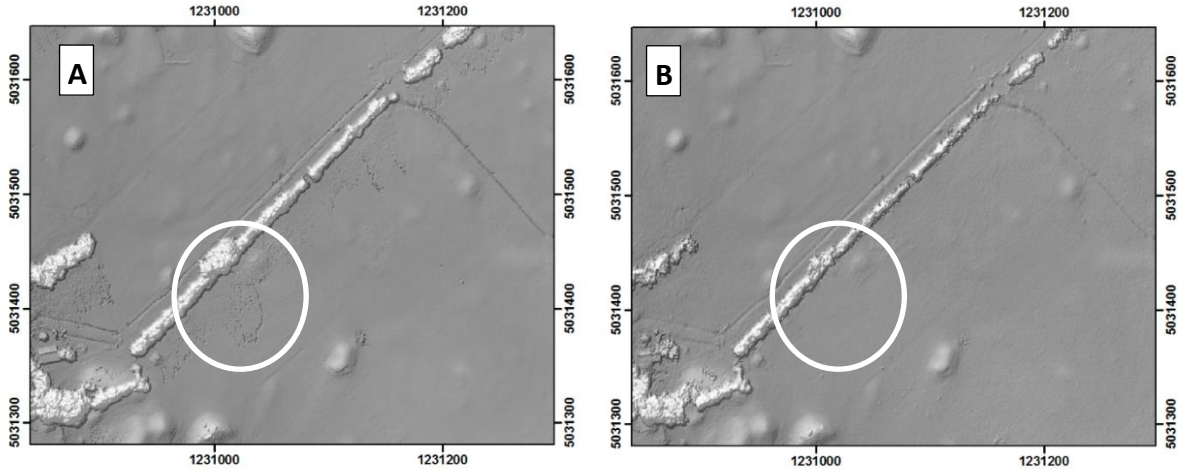

Figure 6.20 Vegetation and regions of shade represented through the hillshade for the A) UAS-SfM survey and B) reach wide survey. *Projection: NZTM 2000*

In this Figure, significant noise is evident in the UAS survey in areas under shadow in the lee of a shelterbelt showing crossing the paddock in a SW-NE direction. Irregular deformation of the flat paddock surface occurs at the margins of the shadows, a pattern that is not present in the helicopter derived model. Both surveys were conducted at similar times of day, so it is likely that such artefacts either stem from differences in the processing algorithms used in TBC and APS, or reflect differences in the viewing angles and image quality.

Characterization of these artefacts demonstrates the power of visual inspection using hillshade models, which illuminates subtle topographic features that are poorly captured by height rendered DEMs. This approach clearly highlights the rough nature of the reconstruction in the wet areas of the channel network, as shown in Figure 6.21. These areas incorporate a ME bias of +1.58 m, reflecting spurious water surface reconstructions in the catchment wide survey, while the UAS survey has the water areas masked from the point cloud, so creating an approximately horizontal water surface, consistent with the bankside water levels. These differences are less evident for minor channels that are not masked from the UAS surface. More broadly, it is also clear that the UAS surface appears to capture finer scale features, despite the equivalent resolution of the DEMs. This reflects the greater depth of information in the raw point cloud, stemming from the higher ground resolution of the imagery due to the lower flying height.

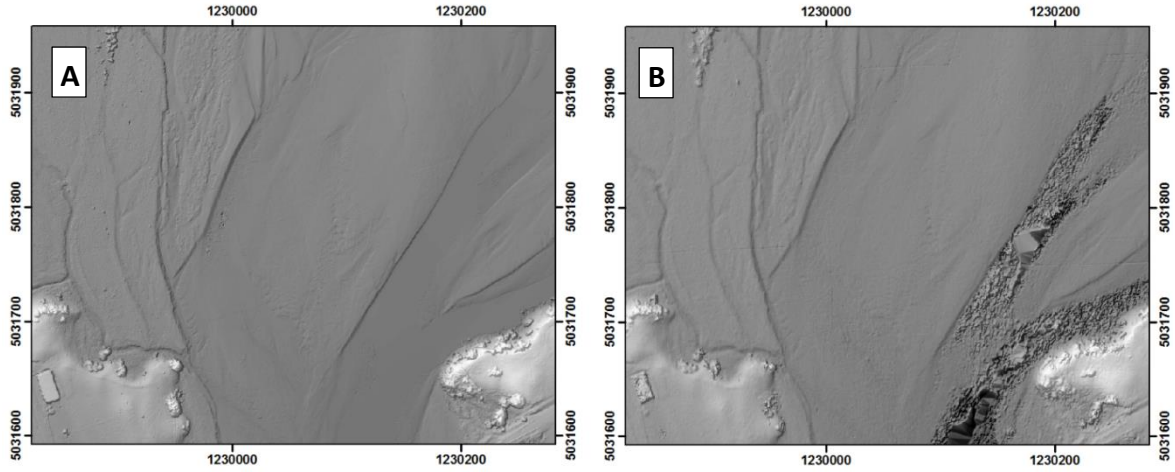

Figure 6.21 Wetted channel represented through the hillshade for A) the UAS-SfM survey and B) the reach wide survey. *Projection: NZTM 2000*

Whilst differences between the wet channels and micro scale bar features are perhaps expected to differ, the hillshade models also indicate differences between buildings the larger scale hillocks themselves as show in Figure 6.22.

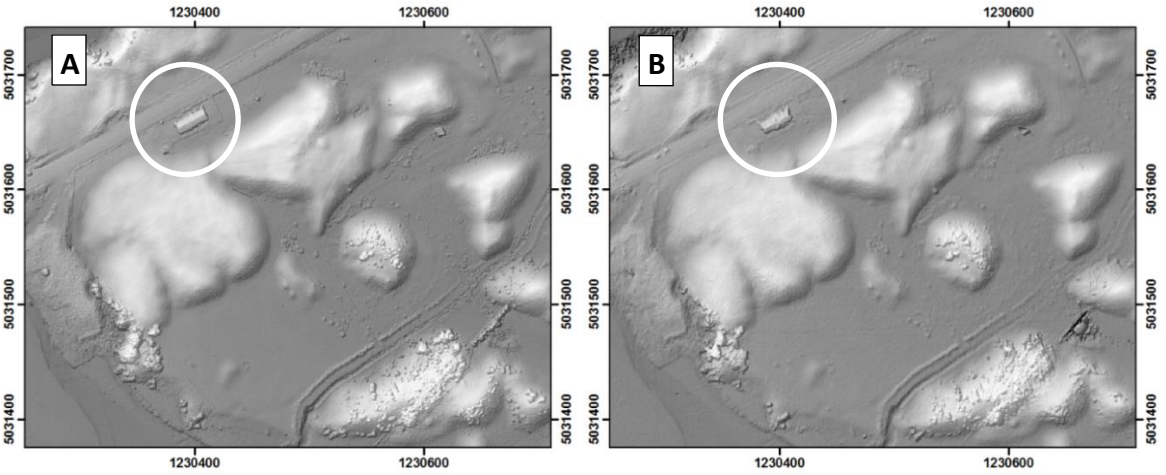

Figure 6.22 Region of variable topography and a man-made structure as seen through the hillshade for A) the UAS-SfM survey and B) the reach wide survey. *Projection: NZTM 2000*

Here again, while the general morphology and position of the hillocks is similar, the UAS survey again captures finer scale features, associated with small terraces, gullies and vegetation patches. The reconstruction of the roadside barn show in the NE corner of Figure 6.22, is much more clearly defined, preserving the sharp edges of the building and roofline in the UAS surface. Such differences relate directly to the resolution of

these features in the native imagery used for each model which is shown for comparison in Figure 6.23. Not only is the barn captured by the UAS at greater spatial resolution, but the enhanced contrast of this image enables precise keypoint identification at a scale that cannot be matched from the higher, helicopter photography. Ultimately, such differences are inevitable given the scale of approach implied in both surveys.

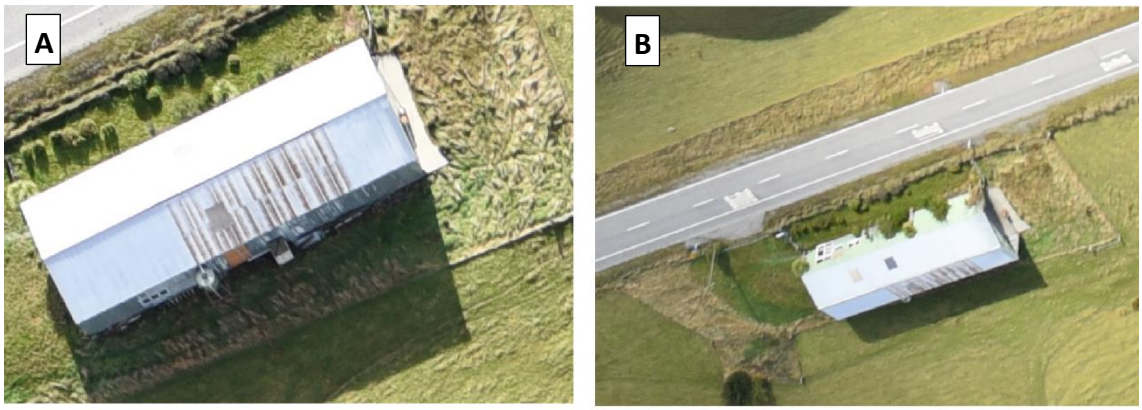

Figure 6.23 Imagery captured from the catchment wide scan of the barn, highlighting the issues faced during the dense pixel based matching stage. The high exposure in (A) results in a lack of features on the barn roof observed in (B).

Whilst the reduced flight height of the UAS-SfM survey has resulted in less noise and an increased fidelity in reconstructing the channel network, vegetation and man-made features, these typically related to random errors that are easily accountable for within an error model based upon surface variance. For the purposes of repeat analysis, the impact of systematic errors is of a much greater importance. It is proposed that impact of systematic errors (pin stripe and relating to GCPs) and imaging and shading artefacts observed within the hill shade DEM render the UAS survey of a lower quality in terms of accuracy than the catchment wide survey – validating the uptake of the 2015 catchment survey protocol as an alternative to standard UAS practices.

6.4. Discussion

6.4.1. **Performance of the Revised Survey Strategy**

This chapter has described the application of a new survey and processing workflow, designed specifically to enable accurate DEM production for wide area SfM surveys. The strategy is aimed directly at mitigating the type of systematic errors identified with the

vertical photography in 2014, but nonetheless reflects the broader challenges associated with conducting surveys over this large scale in a hostile, dynamic environment. A brief summary of the advantages gleamed from each of the key revisions to the survey strategy is given in Table 6.9.

Table 6.9 Broad impacts of survey strategies adopted within the 2015 reconstruction and potential for future work at equivalent scales.

Survey Improvement	Environmental/Logistical Challenge solved	Evaluation for catchment scale application
Convergent Imagery	• ***Large Survey Extent*** • ***Labile Environment*** • ***Complex topography***	The use of convergent imagery is a requirement for future studies over such scales. This camera network geometry provides a stronger basis of accurate camera calibration and pose estimation in the initial bundle adjustment and so relaxes the dependence of subsequent optimization using ground control as additional constraints. This is particularly important for the characteristically low density ground control networks used in these large area surveys and in environments where placement of GCPs is heavily constrained.
Camera GPS	• ***Large Survey Extent*** • *Labile Environment* • ***Computational Resources***	The introduction of a priori camera position information through an on board camera GPS also supports improved estimation of the initial bundle adjustment. Rather than being used as external constraints directly, the camera position data is used to provide a targeted comparison of images during correlation, which enables faster computation times and a more restricted set of likely matches to be retained. However, while useful, significant issues were encountered through the use of low grade GNSS receivers and the experience here may suggest that investment in measurements of platform position with a dual frequency GNSS receiver could be worthwhile. Such data could be post-processed to provide PPK positions at 10 Hz using standard receivers, greatly improving the quality of a priori camera pose.
Imagery format	• ***Large Survey Extent*** • ***Financial Constraints*** • ***Computational Resources***	Capturing images in compressed (JPEG) format is recommended for future studies over such scales. The resolution of imagery and preservation of the majority of keypoints appears to be sufficient to provide reliable topography at the decimetre scale shown here. Reduced file sizes enable higher capture rates, once more enabling increases in flight speeds and further reduce the time required in the initial bundle adjustment.

6.4.2. **Improved surface accuracy: Convergent Imagery and Camera GPS**

In summary, a nadir aft-nadir camera system has been successfully deployed to capture a convergent camera network that incorporates limited information on camera pose through an on-board GPS. The resulting model has significantly reduced ICP residuals (MAE = 0.23 m) compared to 2014 and much more moderate systematic bias and limited effects of GCP placement on the optimization process.

Initial concerns relating to higher surface noise associated with convergent imagery have been shown to diminish through the densification process. This reflects the use of the pairwise SGM algorithm, which uses the sparse cloud only to seed the matching of stereo-pairs. As such, noise in the sparse cloud does not to propagate into the final surface.

The use of camera GPS has proved difficult to implement, requiring extensive correction to enable successful reconstruction within APS. However, while the image location estimates are prone to errors associated with time averaging leading to drift, linear correction equations could be derived to correct for this by comparing the *a priori* position estimate with a posterior solution based on the GCP data. The revised positions enabled the whole study region to be reconstructed quickly as a single block, without major areas of data loss of singularities in the point cloud. Without these *a priori* estimates of camera position, the scale of the model would have required iterative construction as a series of smaller blocks, each requiring optimization and the fusing. This is likely to have led to problems associated with matching the peripheral areas, creating further systematic errors in the model.

6.4.3. Interaction between GCP distribution, focal length and radial distortion estimation

Attributing the improvements in model quality to either the combination of convergent imagery and/or the camera GPS directly is difficult. The preceding chapters have identified the impact that camera pose and convergent imagery have upon the internal and external calibration within the bundle adjustment. Whilst work upon the flume suggested a link between focal length estimation and camera height, the simultaneous manner in which the bundle adjustment resolves these parameters renders it difficult to provide an independent assessment. In an attempt to further delineate the impacts

of these parameters the following section considers their interaction with varying GCP distributions for a 5 km reach based at the Hillocks.

A series of experiments have been created by altering the GCP network of this reach to provide a range of scenarios relating to distribution and density of the GCP network. These experiments were designed to assess the interaction between the focal length and radial distortion parameters with varying GCP networks, to determine if either is more reliant than the other. A total of 906 cameras were used to build a sparse point cloud under high alignment settings comprising 714,294 points, which was then georegistered and optimized using four different combinations of GCP layouts, see Figure 6.22. A baseline (hereafter, *reference*) model was created which used all available (n=17) GCPs (Figure 6.24a), against which the models optimized using three differentially degraded GCP networks were compared. The first degraded network used 10 GCPs but retained a broadly similar distribution to the full network of the 17 used in the reference model (hereafter termed the *sparse* network, Figure 6.24b). The second network used just 7 GCPs, but this time clustered in the southern margin of the block (hereafter termed *uneven*, Figure 6.24c). The final network used just four GCPs positioned in the corners of the block (hereafter termed *peripheral*, Figure 6.24d).

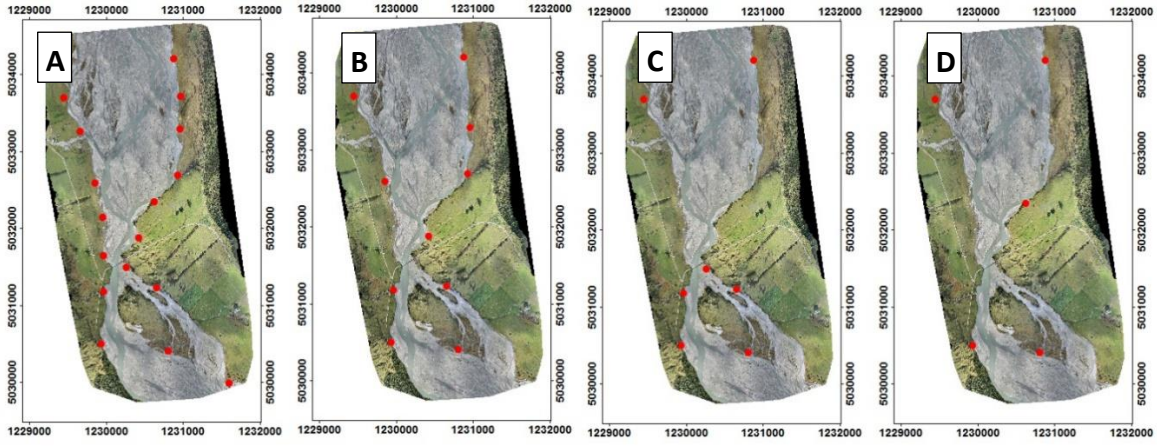

Figure 6.24 GCP layouts for parameter interaction experiment: A) The full network; B) Sparse network; C) Uneven network; and D) the Peripheral network. *Projection: NZTM 2000*

The resulting focal length and K1 parameters following optimization were then compared against those from the initial *a priori* bundle adjustment and for each degraded network against the full reference model (Table 6.10).

Table 6.10 Percentage change in camera parameters of each scenario when compared to the baseline scan.

GCP Layout	% Change difference in optimized		% Change in parameter following	
	Focal Length	K1	Focal Length	K1
Reference	n/a	n/a	0.08	0.83
Sparse (1)	-0.05	-0.08	0.04	0.91
Uneven (2)	-0.05	0.05	0.04	0.72
Peripheral (3)	-0.05	-0.19	0.04	0.91

The change in focal length between the reference alignments is constant across all GCP distributions at a value of -0.05%. In contrast the K1 parameter is far more varied, with the greatest change of -0.19% observed under the peripheral distribution of GCPs. There is also a change in direction of change when considering the uneven distribution as the parameter adjusts positively by 0.05%. The nature of change experienced by the K1 parameter under each varied GCP scenario would indicate that the GCP network is more closely linked to the radial distortion refinement. Despite this the relative degree of change between all scenarios and the Reference GCP distribution are small, suggesting that SfM reconstructions are robust to small variations in GCP distribution which is beneficial for the purposes of conducting repeat experiments.

The degree of change in parameters during the optimization process is significantly different between radial distortion and focal length parameters. The focal length undergoes minimal change across all scenarios. As expected, assuming a greater reliance upon the GCP network, the radial distortion parameter varies more significantly than focal length during the optimization. There is also a greater range between scenarios tested with 0.72% (Uneven GCP distribution) to 0.91% (Peripheral GCP distribution), representing a 26% change between the two. However, the maximum deviation from the Reference setting is just 13%.

Interestingly, results show that the way that the radial distortion and focal length parameters interact with the GCPs is different. In this scenario it would appear that the GCP distribution has a more significant impact on the subsequent adjustment of the radial distortion parameter than the focal length, which itself has been proposed to be more dependent upon camera pose (see Chapter 5).

6.5. Conclusions

This chapter examined a revised survey strategy for wide-area SfM surveys following designs developed from the laboratory experiments in Chapter 5 and the experiences associated with the initial survey in 2014. The key conclusions from the chapter are as follows:

- The use of a convergent camera network alongside camera GPS result in an improvement in overall model accuracy and reduce the appearance of systematic errors. Check point errors show marked improvement from 2014, with MAE reduced from 0.44 m to 0.23 m, MaxE reduced from 2.67 m to just 0.65 m. The effects of optimization on the resulting point cloud are more limited as a consequence of the improvements in the initial bundle adjustment and this limits the distribution of ground control required.
- Comparisons with a benchmark high resolution UAS survey show a generally good fit between these two models, with over 68% of the coincident areas matching by within 0.25 m vertically. There are no major systematic errors in the helicopter survey revealed by this analysis, although the UAS survey does exhibit artefacts associated with overlapping flight lines and shadows. The impact of pixel resolution is significant however, with the UAS survey capable of resolving finer scale topographic features.
- Convergent image networks appear to generate more noise in the sparse cloud generated by the initial bundle adjustment. However, the SGM algorithm used in APS appears to minimize this effect in the production of the dense cloud failing to propagate through to the dense reconstruction.
- Oscillating, wave-like systematic distortion is still evident at the scale of the whole model however has been dampened in wave height. Experiments in chapter 5 (Figure 5.15) have shown that this pattern relates to both focal length and radial distortion parameters.

6.6. Future Work

The aim of this thesis is twofold; to both evaluate the application of SfM for regional (catchment scale) surveys and to assess the fitness of the models derived for the

purpose of assessing geomorphological change, using the Dart River as a case study. This latter task requires an approach to terrain modelling that meets limited requirements of reproducibility and clearly understood errors. Together these conditions enable the definition of an appropriate minimum level of detection that can be used to separate changes that are likely to be natural or 'real' from those that arise due to data errors. The results presented here suggest that the 2015 survey strategy has the potential to provide such consistent models over multiple epochs, facilitating the use of DoDs to analyse geomorphological change with a vertical resolution that is sub-meter in scale. Unfortunately the 2014 survey may not be a useful basis for such comparison. This model incorporates large systematic errors relating to the inferior survey strategy and limitations in ground control placement. Future work is therefore required to determine whether these two surveys can be used together to assess changes on the Dart River over the 2014-2015 period. This will require either: (1) further efforts to eliminate systematic bias from the 2014 survey; or (2) attempts to correlate the two models so that significant changes can be identified by their departure from a reference translation model.

7. Chapter 7: Undertaking Geomorphological Change Detection

Abstract

Previous research detailed in Chapter 6 has demonstrated the use of Structure-from-Motion (SfM) to derive catchment scale DEMs with reduced random and systematic errors (ICP MAE = 0.23 m). For the purposes of change detection it is important to assess the reproducibility of the methods used and the likely variation in repeat surveys in terms of local surface 'noise' and systematic errors at the scale of the whole photogrammetric block.

Localized random noise in SfM models relates principally to the precision and repeatability of keypoint identification and correlation, which is affected by surface texture, contrast and lighting conditions. Such effects can be treated as independent of the systematic errors that reflect inaccuracies in the modelled scene structure in terms of the solved interior and exterior geometry that may average out such local effects. An assessment of the magnitude of surface noise can be gained by comparing identical repeated reconstructions based on the same set of images and exterior constraints. This approach is similar to that used elsewhere to identify a minimum Level-of-Detection for DEM differencing (e.g., Brasington et al., 2003; Lane et al., 2003). The analysis of variability in such repeat models, calculated pixel-by-pixel, can be usefully expressed in terms of simple measures of variance, such as the standard deviation or a statistical confidence interval (making assumptions about the distribution of variability). This approach was trialled here using a series of experiments and used to identify surface uncertainties for different types of surface cover (dry bed, wet bed, vegetated). These surface uncertainties were then used to define a level of detectable change, measured at the 95% confidence interval assuming a similar structure of errors in both the before and after surfaces. This analysis was used to characterise the detectable change for different combinations of surfaces in the two epochs, i.e., dry-dry and wet-wet, which was found to vary from as low as 0.28 m to 2.02 m at the 95% CI.

Local random errors are relatively easy to quantify by comparison with the systematic error that arises from inaccurate camera calibration and the distribution of ground control. Due to the nature of the bundle adjustment which incorporates the ground

control to constrain the surface geometry, it is impractical to provide an empirical estimate of model distortions based upon GCP residuals. Furthermore, differences between the survey strategies in 2014 and 2015 imply that a standardised methodology to correct systematic errors will not be applicable to both models, and the pattern of errors is likely to reflect the unique placement of GCPs for each survey. In response, an empirical approach is adopted in an attempt to correlate the 2014 and 2015 surveys through the identification of regions that are assumed to be stable between the two epochs. This approach is trialled for a 17 km^2 reach with relatively plentiful and well constrained control. Correlation of stable areas provides a means to correlate the 2014 model to the 2015 model. Applying the localized treatment of errors to define a LoD then provided a means to determine the geomorphological change between the two years in this experimental reach. The results suggested that the reach experienced similar volumes of erosion and deposition during the period of the order of 6.7×10^5 m^3, representing similar average depths of erosion and deposition of 0.49 m and 0.47 m respectively. While the results provide a first estimate of the pattern of morphodynamics in this period, it should be recognized that major problems exists associated with the correction of the wetted topography. Such work lies beyond the scope of this thesis but clearly reflects an important concern for on-going research.

7.1. Objectives

The use of convergent camera networks with *a priori* information on camera pose (position) can mitigate significant systematic errors in SfM models. The revised survey approach applied in Chapter 6 highlighted the potential of this method to derive decimetre quality topographic data products across wide areas that are of sufficient quality for geomorphological analysis. This is illustrated by check point MAE of 0.23 m and Maximum errors of 0.65 m and significant reduction in the degree of lens correction required during optimization – specifically the K1 radial distortion and focal length parameters.

Whilst the accuracy of the 2015 survey is promising, the production of the model is only the first step in undertaking an analysis into the longitudinal sediment budget of the Dart River. Ultimately, while model accuracy is desirable, accurate change detection simply requires reproducibility which could at least in theory incorporate systematic errors (i.e., relating to datum's) so long as these errors are consistent through time. The evaluation of DEM quality in this context therefore requires the assessment of two differing aspects of model performance: a) internal consistency of model precision; and b) external measures of model accuracy. These characteristics are summarised in Table 7.1.

Table 7.1. A summary of consistency concerns that may impact upon the ability of SfM models to be used to perform catchment scale geomorphological change analysis, relating to accuracy and precision.

	Type of inconsistency/source	**Impact of inconsistency**
Internal	Errors resultant from variations in repeatability of an identical scene, e.g. surface variations resulting in inconsistent keypoint matching.	Local regions of noise and potential holes in the reconstruction. Errors are likely to be minimal in comparison to larger systematic issues.
External	Errors resultant from differences in survey techniques between years – most obviously, imaging network and GCP distribution.	Significant systematic errors as surveys may vary in quality and reliance upon ground control. Particularly noticeable in areas of inconsistent GCP distribution across surveys.

A host of 'internal' factors may impact the quality of reconstructions 'within' a survey, such as local variations in contrast, lighting and surface types. These are identified as

sources of error that are likely to manifest regionally, linked to the ability to provide reliable keypoints and correspondence. Within this study there is an absence of sufficient checkpoint data across regions that may be prone to internal variations, ultimately limiting our ability to quantify these uncertainties based on traditional residual based approaches (Brasington et al., 2013; Lane et al., 2003). A more novel approach would be that developed by James et al. (2017) to provide 3D precision maps describing the 3D photogrammetric and georeferencing uncertainty through a Monte Carlo based analysis. This process is however reliant on repeated bundle adjustments, which remains impractical for a survey of this scale.

Therefore the primary source of internal consistency that this chapter will focus upon is software variability. This is to say, under consistent exterior constraints such as imagery, GCPs and alignment settings, how precise is the resulting cloud reconstructed by APS. This internal error is a product of inconsistencies in both the keypoint identification and matching (alignment) and the dense SGM algorithm. It is inherently difficult to separate these two sources of error as the keypoint identification propagates through the bundle adjustment to the seeding of imagery during the densification process (see section 3.4). Comparing consistent reconstructions permits a pixel-by-pixel analysis of variance enabling the spatial characterisation of local error recorded across varying surface types.

The quantification of local random errors through an assessment of surface variance within DEMs may be used to derive a spatial model to inform change detection analysis studies. Recent approaches to DEM analysis consist of applying spatially variable fuzzy logic functions (Wheaton et al., 2010) or user defined spatially variable error models (Milan et al., 2007). However, the representation of software variance may be achieved quite simply through the establishment of a minLOD (Brasington et al., 2003; Lane et al., 2003) which determines a threshold beyond which change within a model is deemed to be reliable based upon the surface characteristics. An advantage of this approach is the opportunity to apply confidence intervals, an attractive attribute as software repeatability is likely to provide a conservative prediction of the true degree of imprecision within the survey.

Whilst the quantification of internal error, specifically software variability is a relatively easy task, the quantification of systematic error remains impractical. These systematic

errors, characterised in Table 7.1 as external inconsistencies primarily relate to variation in the distribution of ground control and imaging networks. This ultimately refers to the survey accuracy and how variation in this accuracy between surveys may manifest when undertaking a DoD analysis of geomorphological change.

As noted by James et al. (2017), within SfM independent check points (ICPs) remain the best method for determining model accuracy. The results of the 2014 survey as detailed in Chapter 4 identify an ICP MaxE value of 2.7 m. The scale of this value, itself a manifestation of the inherent systematic error within the survey, indicates that the presence of major systematic errors in the DEMs is likely to dominate any analysis of change detection. Given the unique nature of these systematic errors it is also clear that it is unlikely that a universal approach to error mitigation can be developed that could be applied to both surveys. Any such treatment is therefore going to require bespoke modelling.

A crude solution to this problem is to attempt to correlate or 'tie' the surveys together using the 2015 reconstruction as a reference surface, despite the errors it itself incorporates. An empirical approach to this problem involves the identification of areas that are assumed to have changed little between surveys and therefore provide a basis to determine a network of points or areas that can be used to define a 'correction surface'.

The ultimate goal of this chapter is to quantify geomorphological change between 2014 and 2015 on the Dart River. However, to achieve this, a series of key objectives are necessary. These include:

Objective 1: Establish a crude error model based upon random error (precision) relating to software variability to provide a range of minLODs experienced within the study.

Objective 2: Using the results of objective 1, provide an assessment of the inherent accuracy of a simple DoD through progressive thresholding of a minLOD

Objective 3: Mitigate systematic errors shown to persist through the application of objective 2 through empirically tying the 2014 and 2015 survey together.

7.2. Quantifying random error

While local, 'random' errors intrinsic to the SfM pipeline are likely to be of minor significance compared to the systematic bias, it is a useful first step to evaluate what magnitude of vertical changes could be identified beyond the natural software variability. Random errors specific to software variability are typically impacted by keypoint identification and matching along with the densification process, typically impacted upon by surface type/homogeneity and lighting conditions. A number of previous studies of DEM differencing have used Fuzzy Inference systems to characterise spatial uncertainty based upon a series of user defined metrics (Wheaton et al., 2010). This study is however limited by the scale and complexity of the catchment. The black box nature of the APS software furthermore, precludes ready analysis of the precise source of errors.

The derivation of reliable results from a DoD presupposes a level of consistency or reproducibility of the survey under controlled conditions. To perform such an assessment, reconstructions were performed using identical imagery and the same prior constraints on the bundle adjustment to best represent the 'ideal' repeat survey conditions. This 'ideal' repeat is designed to test the impact of software variability, specifically in relation to the ability to consistently identify photogrammetricly challenging surfaces such as vegetation and water (Lane et al., 2000; Fonstad et al., 2013). As noted previously, the final dense output within APS may be subject to imprecisions accumulated within alignment and densification stages.

7.2.1. **Establishing Uncertainty**

The uncertainty in an ensemble of independent DEMs can be usefully characterized in terms of the pixel-by-pixel variance of elevation estimates for each point in space, i.e., z(x,y). This can be expressed as the local standard deviation of elevation and then either

aggregated across the whole surface to provide a global spatial average, or segmented into regional averages reflecting characteristics surfaces (wet, dry, vegetated etc.), or model contexts (overlap density, proximity to GCPs, shadow etc.).

In order to facilitate the comparison of multiple reconstructions, a small area of the Hillocks reach was chosen. An initial reconstruction was carried out within APS using 140 cameras under the High Reference alignment setting, with a keypoint limit of 40,000 and a limitless tiepoint limit. The initial cloud consisted of 496,000 points with a mean reprojection error of 0.22 (0.74 pix), whilst the dense cloud consisted of ~1.4 x 10^7 points. The reconstruction was georeferenced from five GCPs and optimized, to result in a 3D GCP MAE of 0.003 m. This region incorporates a wide range of land covers and has a robust network of ground control to help reduce the impact of systematic errors. An overview of the workflow employed to derive equivalent repeat reconstructions is detailed in Figure 7.1.

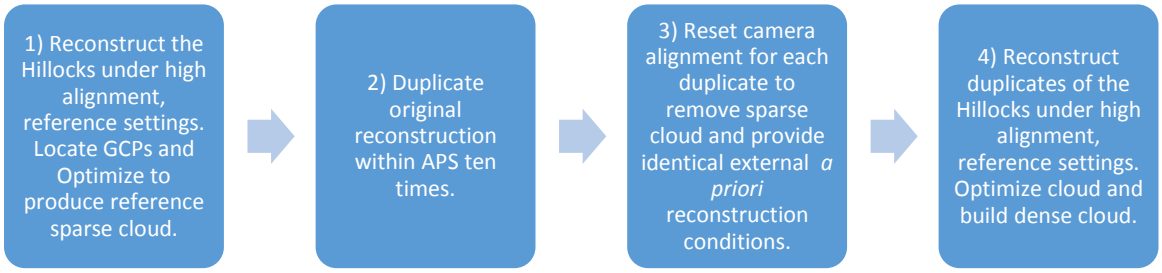

Figure 7.1 Workflow for the generation of models to assess the internal consistency of SfM terrain products.

The initial reconstruction (step 1) detailed prior, including sparse cloud and geo-referenced imagery was duplicated within APS ten times (step 2) to create a series of new chunks. The reset alignment command was used on each duplicate (step 3) to remove the sparse cloud and camera calibration information. This resulted in a set of pre-alignment conditions across all duplicates that were constant including the identification and location of GCPs, essentially reducing the only variance to that introduced by APS. This process varies from the approach taken previously in this thesis by introducing the extra constraint of ground control into the initial alignment phase of the bundle adjustment. This is done to enable consistency in scan location and the

optimization phase between reconstructions which may otherwise be impacted by the manual mislocation of GCPs within APS across repeats.

Each duplicate reconstruction was aligned under the High Reference alignment settings, optimized and densified using the Medium setting with aggressive filtering (step 4). The exported point clouds followed previous procedures, through decimation within Topcat (Brasington et al., 2012) and the generation of a 0.5 m resolution DEM. The Cell Statistics tool (ArcGIS, 2017) was used to provide spatial representations of the SD across all 10 repeat reconstructions and to determine a mean value of SD per surface layer.

7.2.2. **Results**

To verify the precision of duplicate reconstructions Table 7.2 provides the key descriptors for each of the repeat models constructed (n = 10). The inclusion of the K1 parameter and Focal length is to provide a means to assess the geometric consistency and safe-guard against any irregular reconstructions that may impact to skew the results of a variance analysis.

Table 7.2 Key metrics obtained from repeat reconstructions of the Hillocks region aimed at testing the internal (software) consistency of SfM.

Scan	Number of points (Sparse Cloud)	Number of points (Dense Cloud)	GCP error (m)	Camera error (m)	K1 Parameter	Focal Length (pix)
1	2.58×10^5	1.36×10^7	2.36×10^{-3}	1.78	-0.119536	4772.64
2	2.62×10^5	1.35×10^7	2.54×10^{-3}	1.80	-0.119553	4773.38
3	2.59×10^5	1.36×10^7	2.37×10^{-3}	1.85	-0.119599	4773.09
4	2.58×10^5	1.31×10^7	2.58×10^{-3}	1.79	-0.119577	4773.73
5	2.59×10^5	1.32×10^7	2.55×10^{-3}	1.75	-0.119640	4773.85
6	2.63×10^5	1.29×10^7	2.45×10^{-3}	1.83	-0.119647	4773.67
7	2.58×10^5	1.32×10^7	2.45×10^{-3}	1.82	-0.119649	4774.09
8	2.58×10^5	1.29×10^7	2.30×10^{-3}	1.78	-0.119541	4773.26
9	2.61×10^5	1.34×10^7	2.43×10^{-3}	1.82	-0.119587	4773.15
10	2.59×10^5	1.37×10^7	2.52×10^{-3}	1.75	-0.119550	4773.83
SD	$\mathbf{1.67 \times 10^3}$	$\mathbf{3.22 \times 10^5}$	$\mathbf{1.18 \times 10^{-4}}$	**0.03**	**0.0000577**	**0.92**

The results in Table 7.2 exhibit minimal variation across most metrics, indicating a high level of consistency in the reconstructed geometry. The camera error, K1 and focal length parameters, which have been identified as crucial in controlling geometric accuracy display only minor discrepancies. Interestingly, the number of tiepoints varies

in the order of thousands between reconstructions despite using identical imagery. This is significant as the identification of tiepoints is the first stage of implementing a bundle adjustment and propagates down the pipeline by informing the seeding order of images during the densification process.

The absence of systematic errors between these reconstructions is illustrated by plotting the local SD of elevation shown Figure 7.2.

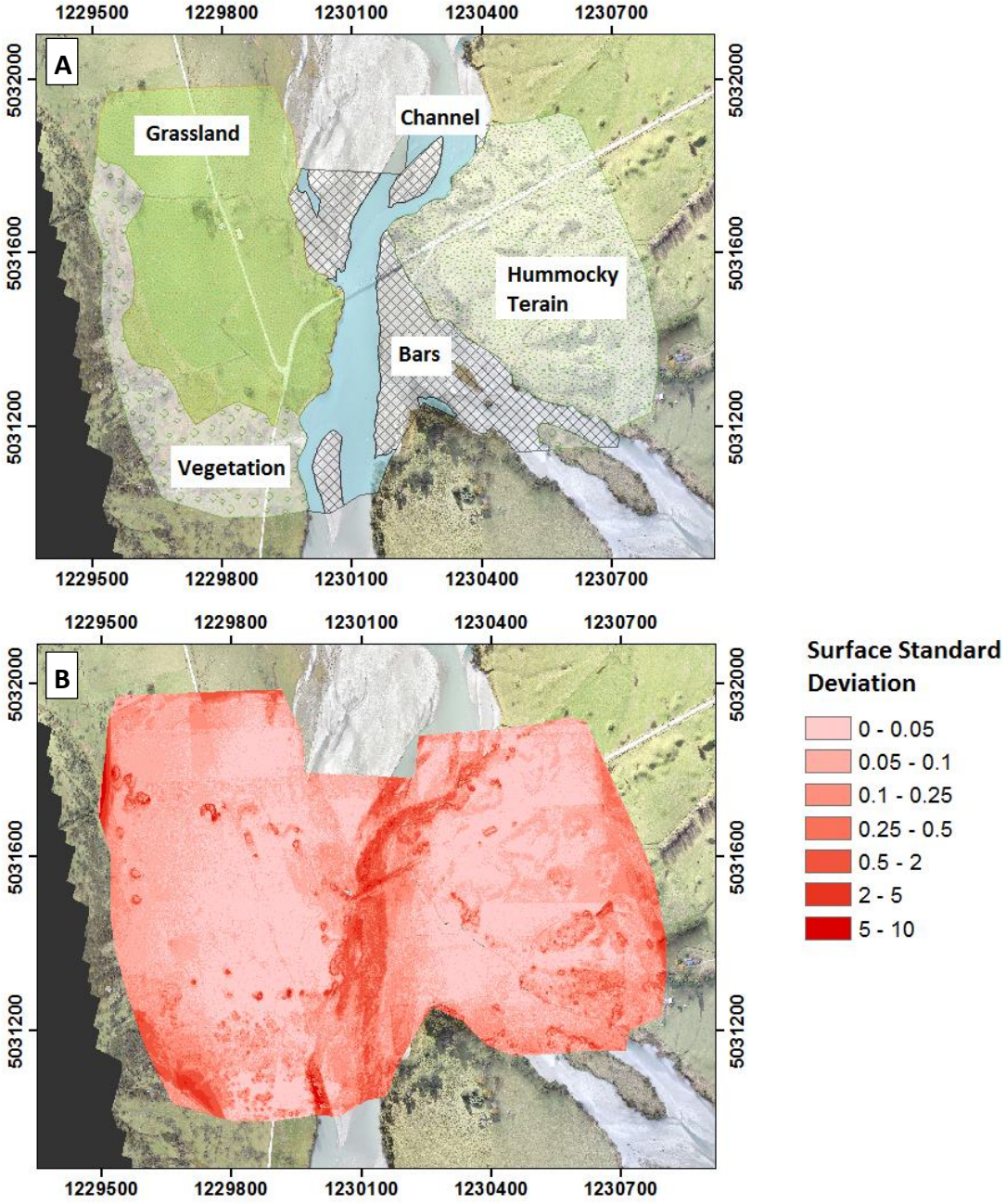

Figure 7.2 (A) Segregation of land types for the repeat surveys at the Hillocks and (B) a spatial representation of SD as derived by ArcMap cell statistics. *Projection: NZTM 2000*

The SD of elevation shown in Figure 7.2 will not account the presence of structural, systematic errors, but does illustrate how the local precision of the elevation varies across the block. As such by comparison with the orthophoto, much of this variability is closely tied to surface cover, with very high variance associated with the wet channels. This is to be expected as the rough, deformable water surface is unlikely to preserve keypoint geometries that are identifiable in multiple images, taken across a wide range of times. While, the vertical uncertainty is dominated by this surface cover, there are further secondary effects associated with increased variance in densely vegetated areas and even flat but grassy regions. These results suggest a significant component of the noise in the resulting DEMs relates to the effect of surface cover on the quality of pixel correlation during the SGM depth mapping. In order to account for these effects, the orthophoto was classified manually into five surface classes: exposed gravel bars; inundated channels; short grass; hummocky terrain with tussock grass; and tall vegetation as shown in Figure 7.2a. The spatially averaged standard deviation of pixel elevations from the set of 10 models was then determined for each class and the results presented in Table 7.3.

Table 7.3 Land class specific SD to provide a value for equation 7.1, with SD of ASD to provide an indication of consistency within class.

	Exposed Gravel Bars	**Inundated Channels**	**Short Grass**	**Hummocky Terrain with Tussock Grass**	**Tall Vegetation**
Average SD	0.10	0.73	0.10	0.13	0.36
SD of ASD	0.13	1.08	0.32	0.30	0.83

The average SD across surface classes is highly varied, with significant variance as expected present for inundated channels and to a lesser extent tall vegetation. The high channel SD is attributed to spurious tiepoints and the inability of the SGM algorithm to eliminate these as a result of the turbid homogenous and moving water surface. This is further exacerbated under the flight height and speed conditions required to perform a catchment scale survey. This has been seen previously in Chapter 6 when providing a comparison with the UAS-SfM survey. There, the reduced pixel size and sharpness increased the difficulty in reconstructing wetted regions consistently.

The impact of the limitations imposed by the survey strategy upon wetted surfaces are mirrored in those that are vegetated. In these areas the typical standard deviation elevation in replicate surveys is 0.36 m. Although this category exhibits significant within class variability as shown by the spatial standard deviation (ASD) of 0.83 m, which indicates that whilst the mean spread is moderate, the variation in these results has the potential to be significant.

In contrast the remaining classes, exposed gravel bars, short grass and hummocky terrain with tussock grass exhibit relatively small SD values. The ASD values for short grass and hummock terrain are enhanced once more indicating inter class variability, although this is potentially a result of the rather general classification scheme applied. Across both metrics the exposed gravel bars indicate a high level of precision relating to the software repeatability of APS. It is however important to consider how these classes interact with one another to inform any form of change detection, see Table 7.4.

Table 7.4. A matrix of 1 sigma LoD calculated for changes in surface classes based on the variance analysis reported in Table 7.3. Units are (m) of elevation change.

	Exposed Gravel Bars	Inundated Channels	Short Grass	Hummocky Terrain	Tall Vegetation
Exposed Gravel Bars	0.14				
Inundated Channels	0.74	1.03			
Short Grass	0.14	0.74	0.14		
Hummocky Terrain	0.16	0.74	0.16	0.18	
Tall Vegetation	0.37	0.81	0.37	0.38	0.51

The matrix above provides simple values for detection of 1-sigma LoD based on error propagation when considering once class changing to another. For example, the accumulated error when comparing an exposed gravel bar to an inundated channel and thus representing channel change would be 1.03 m. The application of such specific classes over the entirety of the reach is impractical and not in standing with previous approaches. The following shall address the establishment of a minLOD based upon the above analysis, with respects to dry-dry, dry-wet and wet-wet variations.

7.2.3. **Establishing a universal minLoD**

Due to the large and complex scale of this study, the consistency of varying surface classes will be used to inform a minLoD based approach for determining error (Brasington et al., 2003; Lane et al., 2003). When undertaking a DEM-DEM comparison,

the error inherent in both surveys renders volumetric errors doubly sensitive to DEM quality necessitating an analysis of inherent errors (Brasington et al., 2003). Lane et al. (2003) identified this minLoD, through quantifying SDE of RTK-GPS measurements for varying bed types (dry-dry, dry-wet and wet-wet).

Using the results from Table 7.3 we are able to derive a minLoD for Dry-Wet, Dry-Dry and Wet-Wet surfaces. This assumes that a minLoD may be calculated under the assumption the SD values in Table 7.3 obey the same propagation of errors as detailed within Brasington et al. (2003).

$$SD_{SfM} = t\sqrt{SD_{Dry}^2 + SD_{Wet}^2} \quad [7.1]$$

Where SD_{SfM} = Mean SD derived from SfM repeats in section 7.2.3 and *t* is a student t-statistic. Assuming a two tailed test, a value of 1 for *t* represents a 68% confidence limit and 1.96 a 95% confidence limit (Brasington et al., 2003; Wheaton et al., 2010).

Table 7.5 Minimum Level-of-Detection for surface comparisons used to threshold the DoD based on the approaches used by Brasington et al. (2003) and Lane et al. (2003).

DoD Surface	Dry-Dry	Dry-Wet	Wet-Wet
MinLOD at 68% CI (m)	0.14	0.74	1.03
MinLOD at 95% CI (m)	0.28	1.44	2.02

The mean SD for the channel regions results in a minLOD at 95% confidence interval or 2.02 m, with a value at a 68% confidence interval of 1.03 m. Likewise at a 95% CI the minLOD for dry-wet change is 1.44 m. In comparison the mean channel depth for the lower 15 km of the Dart is just 0.82 m, with a 90% percentile of 1.5 m. As a result, the natural variability in wetted channels between two “identical” repeat surveys is significantly greater than the max depth change (dry bar becoming a wet channel or vice versa) possible by 90% of the channels. The values presented in table 7.5 provide a range of thresholds beyond which we are likely to see varying impacts of geomorphological change within a DoD survey.

7.2.4. **Summary**

The establishment of a range of minLoDs has been achieved through analysing the inherent variance within the implementation of both the bundle adjustment and SGM algorithms within APS. This provides a conservative value for precision over varying surface classes and fails to account for the full scope of inconsistencies that could occur in a repeat analysis. For example this approach discounts the imprecision that may arise through varying image sets, ground control registration and lighting conditions. Whilst the precision here is acknowledged to be independent of the global systematic error that persists as a result of variations in GCP distribution and camera networks.

The degree of variation present within tall vegetated and inundated channel surfaces purely through software variability limits their ability to be used within change detection analysis. In the instance of inundated channels, the inherent lack of precision is of an order of magnitude equivalent to the potential maximum change that may be observable. Here, noise, irrespective of systematic bias may be greater than the signal. This is a consequence of the turbidity, homogeneity and motion of these surfaces, which provide constraints on the keypoint quality that is exacerbated under the survey conditions employed in this survey.

In line with the conservative nature of the approach to representing surface imprecision a 95% CI has been chosen to provide a range of vales for dry-dry, dry-wet and wet-wet change. The minLoD for dry-dry change is established at 0.28 m, which despite the methodology employed within this survey, is comparable to previous estimates within digital photogrammetry (Brasington, et al., 2003; Lane et al., 2003; Rumsby et al., 2008).

7.3. Raw and Filtered Models of Geomorphological Change

Chapters 4 and 6 detail the production of a dense, georeferenced point cloud for both the 2014 and 2015 SfM surveys. An initial DoD analysis was undertaken for 2015 – 2014 DEMs generated from these dense clouds within APS, decimated via TopCat and exported for raster generation within ARC. The decimated point cloud was interpolated with the use of a TIN and sampled at a 1 m grid resolution, before the 2014 survey was subtracted from the 2015 to provide a DoD. The processing of DoDs was achieved using

the Geomorphic Change Detection tool (GCD) for ArcGIS (Wheaton et al., 2010). This provides a variety of options to model errors and filter observed changes using different detection strategies. The tool also produces maps of change, elevation change distribution histograms and a statistical decomposition of changes.

The minLoD values presented in table 7.5 provide a sensible range of values from which to undertake a sensitivity analysis of the 2015-2014 DoD to assess the potential survey imprecisions and inaccuracies. A total of seven scenarios are presented in Figure 7.3, relating to the incremental increase of the detection threshold (Raw, 0.25 m, 0.5 m, 0.75 m, 1 m and 2 m). The models presented represent the subtraction of the 2014 surface from the 2015, as such positive changes represent increases in elevation between years and negative represent surface lowering.

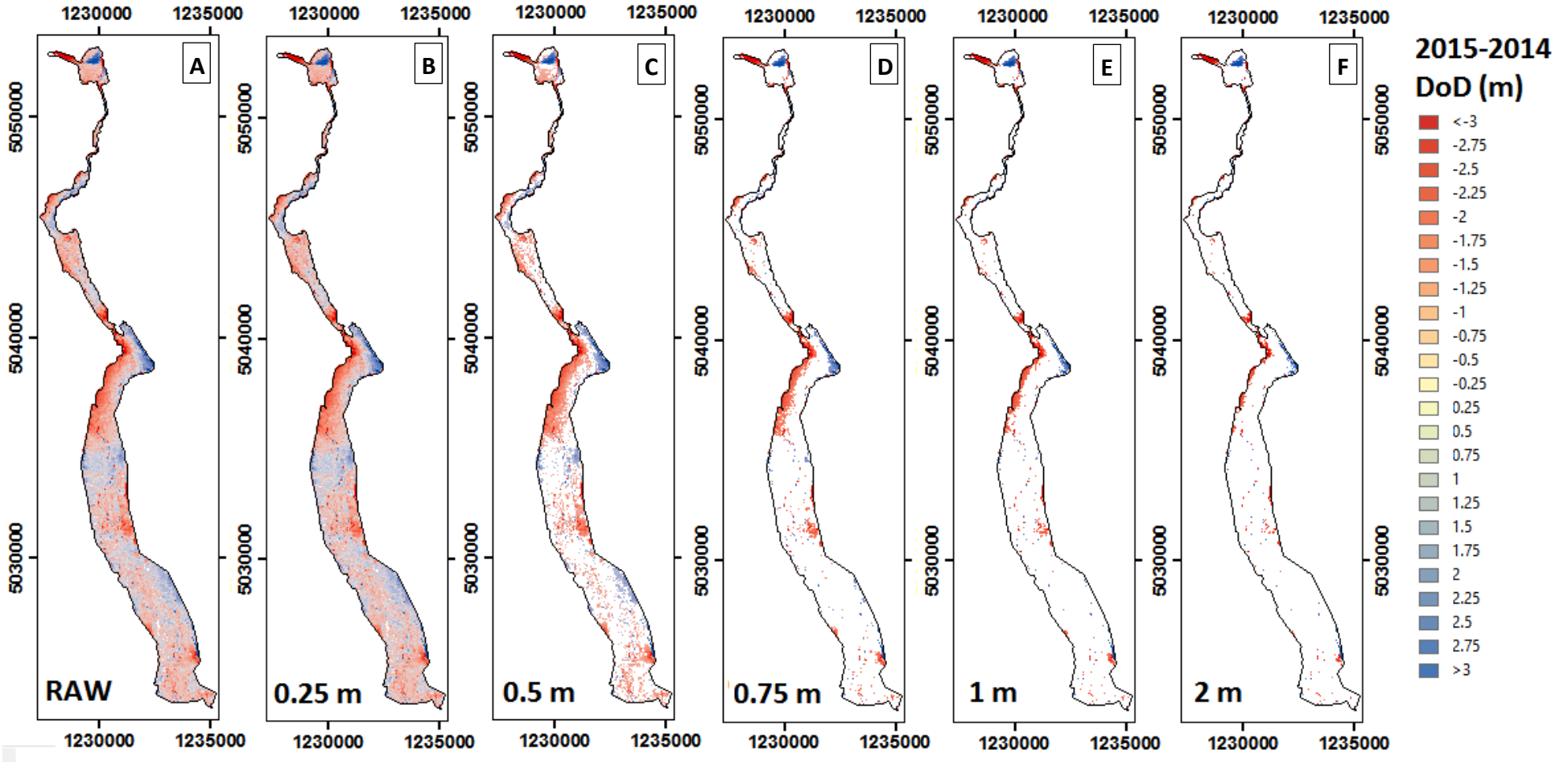

Figure 7.3. Sensitivity analysis of the reach-wide 2015 -2104 DoD, incrementally increasing the minLoD threshold, based upon the values determined in Table 7.5. (A) Raw, (B) 0.25 m, (C) 0.5 m, (D) 0.75 m, (E) 1 m and (F) 2 m. *Projection: NZTM 2000*

The models of geomorphological change shown in Figure 7.3 exhibit evidence of large-scale systematic differences that do not resemble typical fluvial processes. These effects are most clearly evident in the Paradise, Chainman's and Gorge reaches, which are characterized by coherent regions of high magnitude change over 2-5 km areas, significantly bigger than even the largest macro forms in the channel. At Paradise, changes of up to 5 m are identified, both an alternating positive and negative difference on either side of the channel.

By contrast, the pattern of change on the upstream debris fan appears to resemble more characteristic sedimentation and erosion patterns, with surface lowering in the main gully of the landslide, and a dispersed, fan-like pattern of surface increase below it. In this region, the scale of vertical changes associated with this processes is large, with changes of over 10 m lowering the gully and up to 10 m for surface gains along the centreline of the apparent sedimentation. While these elevation changes are clearly high magnitude relative to what might be expected in terms of fluvial erosion and scour, they are not uncharacteristic of what might be expected for gravity driven slope failures and debris flows.

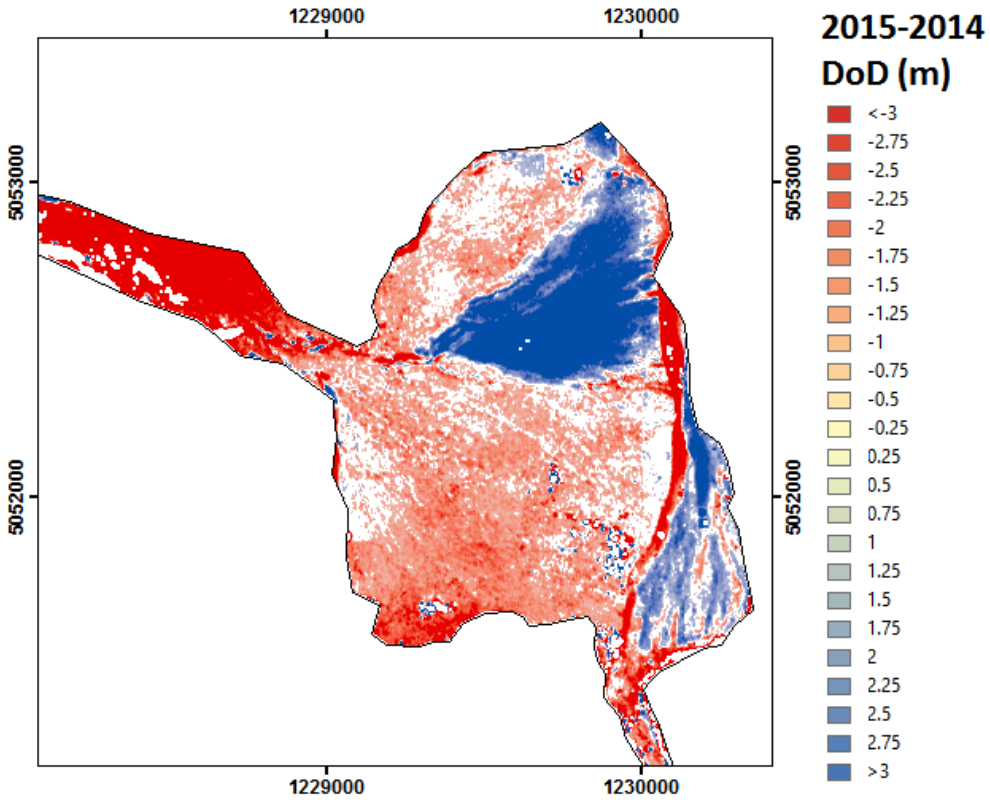

Figure 7.4 The debris fan surface, which provides an example of natural detectable change through accumulation of debris upon the toe. *Projection: NZTM 2000*

Elsewhere, the changes in elevation associated with fluvial processes are difficult to determine against the backdrop of the widespread systematic 'errors'. The analysis of changes in Delta reach, while better constrained in both the 2014 and 2015 DEMs, is confounded by the braiding intensity and dynamism of the channel so that much of the area is either wet in 2014 or 2015 or both and therefore subject to significant surface noise.

The progressive filtering of DoDs reveals the persistence of systematic bias beyond the 1 m minLOD which masks any natural geomorphological change aside from the debris fan. The relative positive and negative changes under each minLOD are presented in table 7.6.

Table 7.6 The relative positive and negative change within the area of detectable change during the reach-wide DoD filtering as seen in Figure 7.3.

MinLOD (m)	Total area of interest with detectable change (%)	Positive change (%)	Negative change (%)
0.25	72%	34%	67%
0.5	47%	30%	70%
0.75	33%	27%	73%
1	24%	25%	75%
1.5	12%	25%	75%
2	6%	28%	72%

There is a constant discrepancy between positive and negative change that is seemingly independent of the minLOD. This lack of equivalency, resulting in a net lowering of the reach is unrepresentative of expected change. The total area of interest with detectable change raises a limitation of the minLOD approach; that is, as the threshold increases to overcome systematic noise, the resulting area useable for analysis is greatly reduced (Wheaton et al., 2010). For example, applying a minLOD value of 2 m provides only change detection data for 6% of the survey extent, missing the majority of natural change,

In the presence of the broader systematic error, it is clear that the statistical metrics of change are unlikely to reflect geomorphological processes effectively. In an attempt to address this, the reduced degree of error within the delta section provides a region suitable for further study aimed at mitigating the differences between 2014 and 2015.

Whilst quantifying the degree of systematic error remains impractical, it is possible to tie the surveys together to enable analysis of morphological change.

7.3.1. Summary

The analysis of geomorphological changes between the 2014 and 2015 DEMs shown above is, with the exception of changes in at Slipstream, ultimately confounded by the scale of the systematic DEM errors. As discussed in Chapter 4 these were found to be related to the combined effects of poorly modelled intrinsic and extrinsic camera geometry and the spatial patterning of GCPs. The complexity of the interaction between estimated components of the 3D structure and its post-hoc optimization using exterior constraints is difficult to unravel given the proprietary algorithms used in APS. Ultimately, such complexity makes the development of a theoretical approach to mitigate these systematic errors intractable.

Quantification of the error structure could, potentially, be tackled empirically, solving the correction of the model inversely. However, the only available regional DEMs for the study area is a 20 m resolution model distributed by LINZ which is constructed from digitized 1:50,000 mapping. There are, as described in previous chapters, small areas covered by high quality datasets, e.g., the TLS model of Slipstream, the UAS model of the Hillocks and the LiDAR dataset of the Glenorchy and the lakeshore. None of these surfaces, however, provides a wide enough coverage to interpret the low frequency oscillatory error structure that was detected through the optimization analysis.

In the absence of a high quality, regional model, the only independent reference data spanning the system-scale are the GCPs and check points, which ultimately provides only a sparse point-based framework for modelling the spatial structure of error. The following section considers a range of approaches that might be used to address this issue.

7.4. Mitigating Systematic Errors

Given the difficulties of objectively 'correcting' each DEM, an alternative approach to facilitate an analysis of the distribution of key geomorphological changes is to correlate the two models. This can be achieved by identifying a network of stable reference areas where the assumed actual surface change is negligible and quantifying the residual difference between the two surfaces. Assuming a representative coverage of matching areas, the residual differences could then be used to 'fit' one DEM against the other by defining an appropriate transformation. Unlike other point cloud registration approaches, such as inter-swath matching for LiDAR (Behan, 2000) or global least-squares cloud matching (Greun and Akca, 2005), here the patterning of residual errors is likely to exhibit considerable complexity over a range of spatial scales, so it is unlikely that any correction could be defined as a simple, global transformation function. Instead, the residual differences would need to be treated locally, the simplest strategy for which is interpolate to generate an empirical correction 'map' which is then applied to translate one model onto the other. This approach assumes that the dominant component of error is vertical and that the surfaces are otherwise matched well in the horizontal.

In order to explore this approach further, a sub-reach incorporating an 11 km section from the Hillocks to the Delta was selected for further analysis (Figure 7.6). This 17 km^2 area was chosen to provide a smaller scale for a proof-of-concept study, but principally, because of the risk that the regional sediment pulse may have already led to significant widespread changes in bed levels upstream which may distort any attempt to correlate the two surface models. This reach was considered to be sufficiently far downstream and wide enough that any elevated sedimentation here would lead to minimal net elevation change.

7.4.1. **DEM Pre-Processing**

Prior to further analysis, all wet areas of the each DEM were first removed. While somewhat drastic, this reflects the significant surface noise identified in these areas that stems from the failure to see through to the bed and the lack of persistent water surface structures that are required to reconstruct the water surface. The scale of the artefacts

that emerge in wet areas of the model, is illustrated by the DoD calculated at the Dart River bridge and shown in Figure 7.5 below.

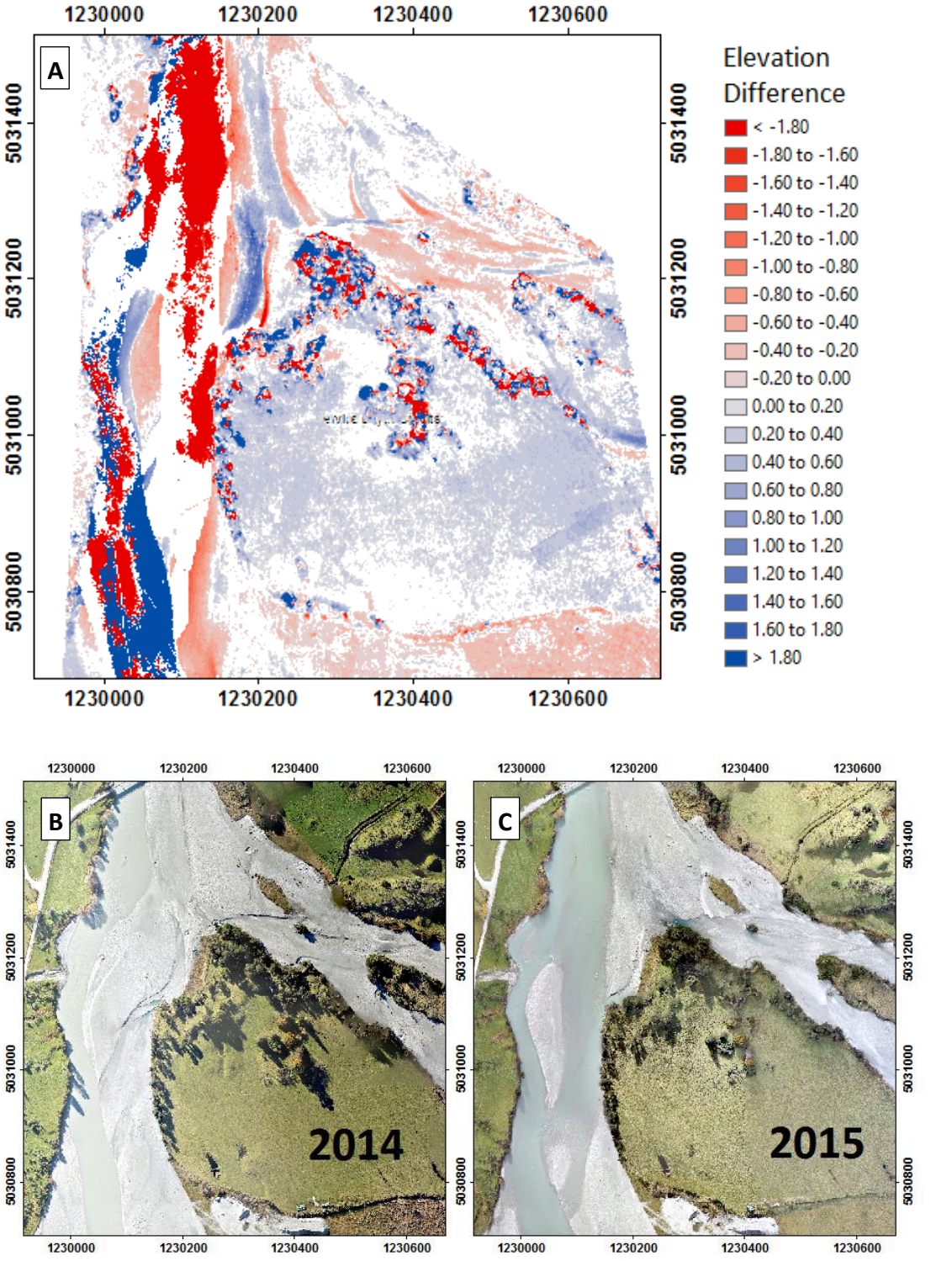

Figure 7.5 Example of the noise introduced into the DoD (A) from significant wetted channels within the Hillocks area which is representative of the entire reach. This noise dominates the detection of actual change between 2014 (B) and 2015 (C). *Projection: NZTM 2000*

Here, large anomalous changes are independent of lateral migration of inundated channels. For example, the significant surface lowering on the true right of the upstream channel would be expected to reflect migration of a wetted channel in 2014 that becomes a dry bed in 2015 survey. Observation of the respective orthophotos in Figure 7.5 show that in reality this is not the case. This is a result of artificially high surface variance (noise on the water surface) dominating any natural change irrespective of the comparison surface.

More generally, an aDcp survey of the whole study reach in 2015 found that mean depths in the major anabranches of the lower Dart were in the region of 0.8 m, with a SD of 0.5 m. Assuming a normal distribution of depths, this would imply that 95% of possible maximum changes, i.e. a wet to dry or dry to wet transition, would fall within a maximum elevation change range of 0.8 + 2(0.5) = 1.8 m. This is broadly consistent with the minLOD associated with dry-wet and wet-wet transitions discussed above (95% CI = 1.44 m or 2.02 m), so that relatively few such changes would be detectable given the likely surface noise.

As such, while representing a major loss of information, it was decided to simplify all subsequent analysis of geomorphological change to exposed bars (dry-dry transitions) only. To achieve this, the wetted channel network was digitized from orthophotographs generated for both the 2014 and 2015 models. These were generated within APS using the default settings and split into 1 km by 1 km blocks for exportation. This manual approach to classification was preferred due to the difficulties of using a spectral based classifier on wet-gravel surfaces that have similar spectral response to turbid water. The resulting vector maps of the channel networks are shown in Figure 7.6 below and used to mask these areas from the DEMs.

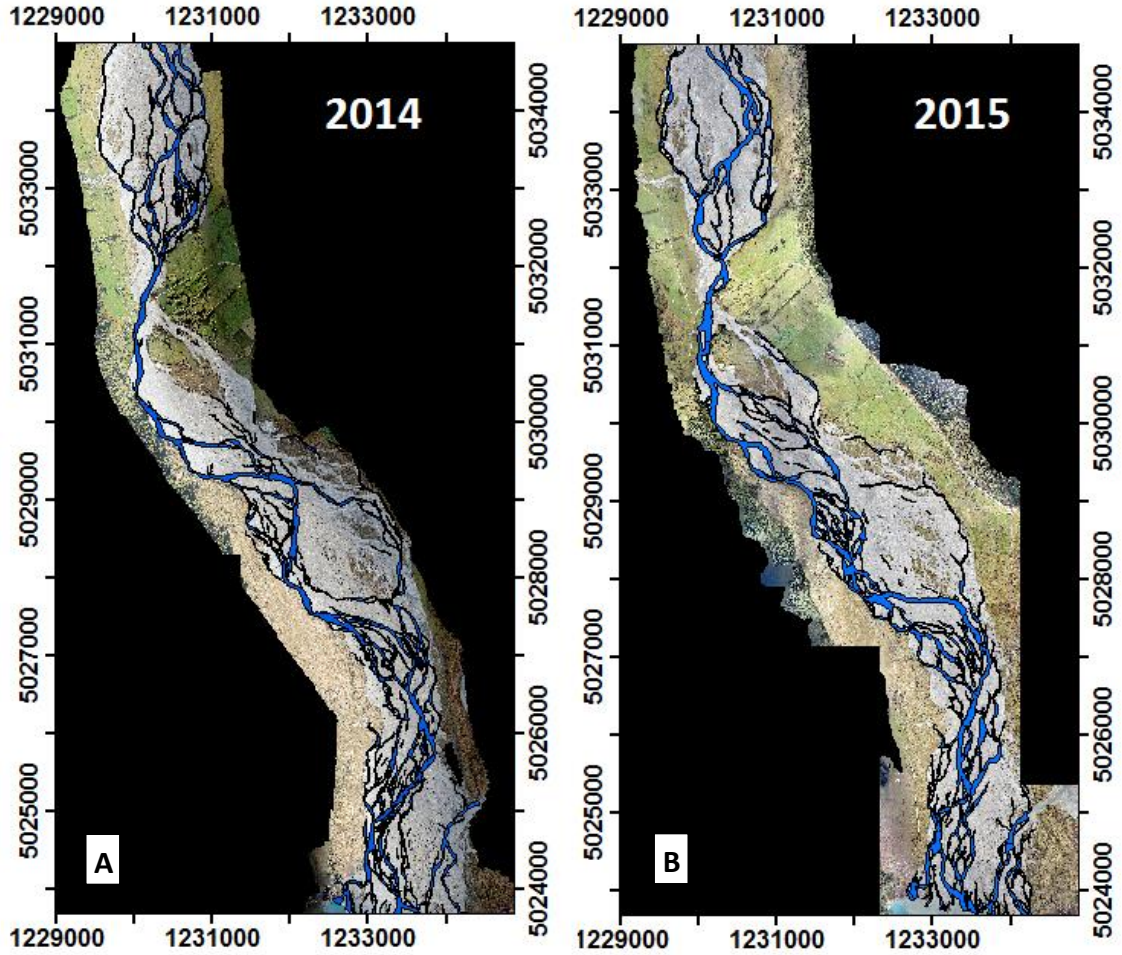

Figure 7.6 Orhtophoto's for both 2014 (A) and 2015 (B) with the respective channel network that was identified and removed from subsequent change detection analysis. *Projection: NZTM 2000*

7.4.2. **Modelling DEM Correlation and Correction**

Two methodologies are proposed for correlating the 2014 and 2015 surveys. The first of these is based upon lateral regions of the model that are assumed to have been undisturbed, while the second approach is based on an analysis of residuals calculated by comparing the 2015 GCP elevations to change values derived from the DoD. For both, approaches, a correction surface was fitted to transform the 2014 surface to the 2015 model. It should be recognized therefore that this approach inevitably preserves the error structure of the 2015 model.

*7.4.2.1. **Stable lateral regions***

This approach involves matching the two models by sampling stable surfaces, similar to the approach used to match strips of LiDAR data (e.g., Felin, 2003). The initial stage in this process is the identification of natural (i.e., not placed targets) that are unlikely to

have been disturbed between the surveys. These include idealized hard surfaces such as roads or roofs, but more likely in the current context, are noisy surfaces such as grassy, high terraces above the active channel. It is important to recognize that an implicit assumption here is that the two models are presumed to be well-matched horizontally, so that only the vertical dimension is adjusted.

Orthophotos were again used to identify a series of stable, lateral regions, on both true right and left banks for each survey. The polygon mask for each survey was digitized and the two sets of stable areas intersected to define the region for correlation (Figure 7.7). Matching involved a sequential process, aiming to generate a low frequency correction model that smoothed the inevitable surface noise associated with pixel matching in the largely vegetated areas.

First, the area of intersection was used to clip the DoD representing the estimated change over the period (2015-2014). Given the presumed stability of this region (i.e., $\Delta Z \sim 0$), changes observed in the DoD reflect the combination of both broad systematic errors and local noise. To eliminate this latter component, but retain the systematic error, a 25 x 25 m low pass filter was used smooth the noise. The resulting surface was then resampled to a 100 m grid and converted to a reduced resolution point cloud.

This process generates two reduced resolution point clouds representing stable areas on either side of the active channel fairway. These were then interpolated laterally across the channel by Delaunay triangulation to create a TIN, which was then resampled at the native 1 m DEM resolution. The assumption in this process is that the systematic errors arise principally in the lateral margins of the model, associated with the placed ground control. Given the absence of constraints in the centre of the active channel, the error function is assumed be averaged laterally across the channel, informed by the deviations from zero measured in the lateral DoD. This new surface therefore now represents the assumed systematic errors in the 2014 model relative to 2015 and can be simple added on a grid cell- by-grid cell basis to either the 2014 DEM or directly to the DoD.

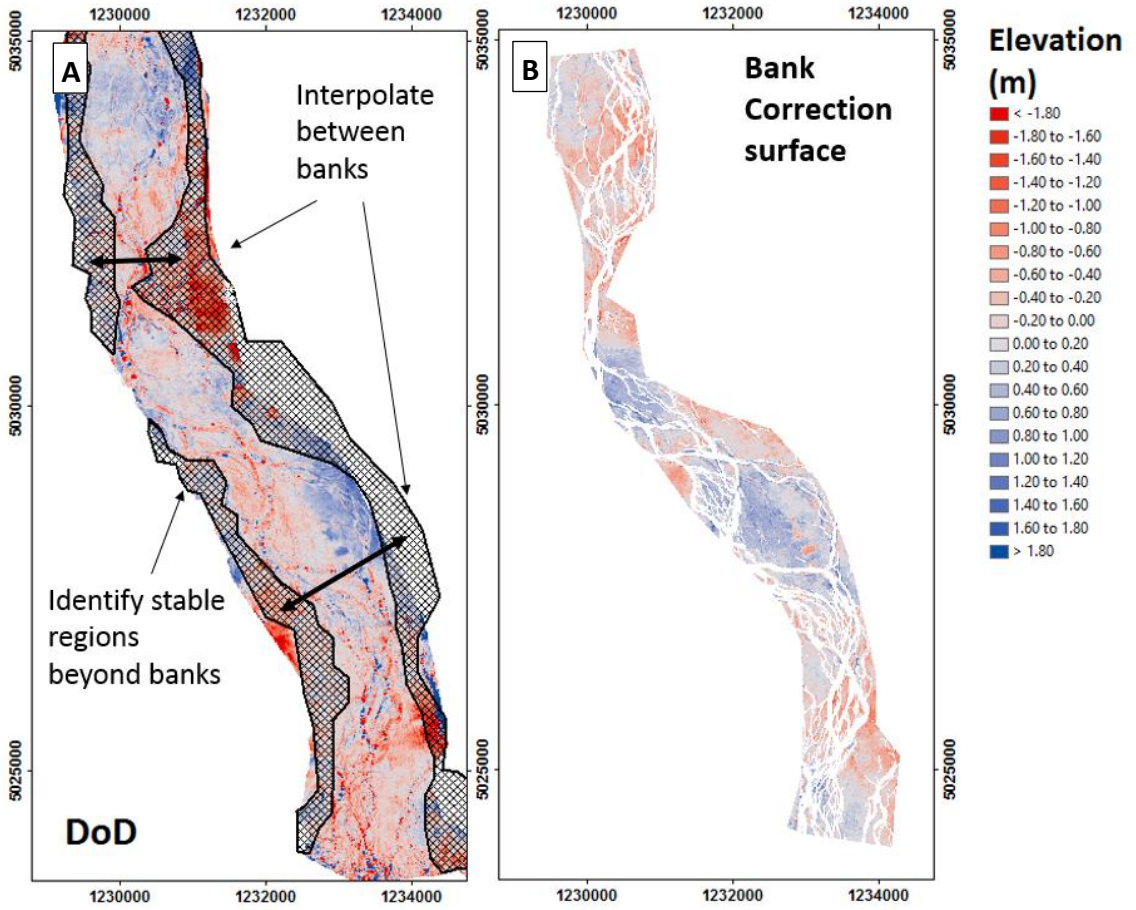

Figure 7.7 A) Schematic diagram of the process applied to derive a correction surface based upon the identification of stable lateral regions. These regions were identified beyond the channel margins within the orthophotos and applied to the initial DoD to identify overlapping regions which may be used to provide spatially consistent areas. B) Subsequent conversion of these regions to a sparse point cloud identifying inconsistencies resulted in the creation of an error mask. *Projection: NZTM 2000*

7.4.2.2. Modelling Systematic Errors using GCP Residuals

The surface correction procedure described above has the advantage that it is informed by spatially contiguous regions presumed to reflect areas of 'no change' along both sides of the channel. This provides a robust basis for interpolation across the wider area of interest, assuming that the lateral structure of error trends smoothly. The approach does, however, suffer from a risk of generalization. Significant local variability exists within the areas presumed to be stable and this is smoothed and generalized by the local and regional interpolation strategy used.

The second data-driven approach to modelling errors relies on the same underpinning philosophy. Here, however, the pattern of systematic error is determined by comparing

the 2015 GCP elevation to the change observed in the DoD (and by inference, the difference between this elevation and the 2014 model). The governing assumption here, is that the GCPs were placed only in areas that were unlikely to be disturbed by high flow events and are consequently, some distance from the active channel. The GCPs should be located in areas with negligible elevation change between the two surveys therefore. Given the comparatively high positional accuracy of the GNSS GCP survey and the tight correlation of the model to these GCPs due to the optimization process, the difference between the GCP elevation and the DoD should be a good estimate of the local systematic error.

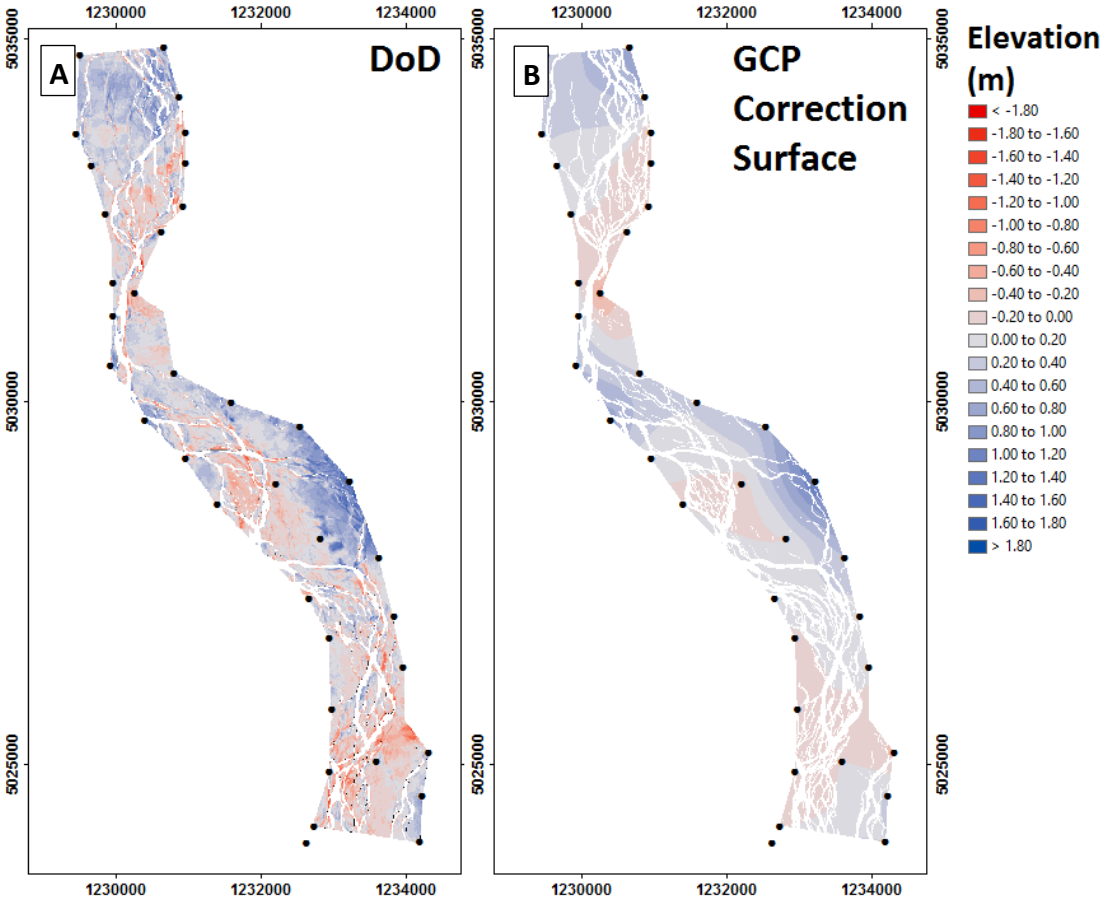

Figure 7.8 A) Schematic diagram of the process to undertake a GCP based correction of the surface. The GCPs are located and measured at each point on the DoD after the application of a low pass filter at 10 m resolution to eliminate noise. B) The resulting sparse cloud may then be interpolated to produce the GCP correction mask observed in (b). *Projection: NZTM 2000*

Implementing the approach described above poses a classic geospatial problem, that being the difficulty of comparing a point measurement (i.e., the GNSS GCP observation) with a spatially integrated measure (i.e., the DEM). Such analysis is subject to local variation in the surface dominating the comparison. There is therefore a need to average the effects of local noise in the DEM before undertaking the point-to-grid comparison. As the GCPs were placed on areas of level flat topography, a sensible strategy here is to smooth DEM surface using a local low-pass filter. Here, a 10 x 10 m low-pass filter was applied, sufficient to account for surface noise, but small enough not to sample major breaks in topography around the GCP.

This analysis is additionally supported by the consistency in the majority of ground control placement over the two years, with attempts made to place targets in similar positions where possible. However, large areas on the true left, as identified previously, suffer from a dearth of control in 2014. To combat this, an additional three control points were identified manually from the orthophotos to attempt to better characterise the structure of apparent deformation. With these additional points, a total of 37 point measurements are available to quantify and model the surface correction.

To define a full correction surface, the 2015-2014 residual at each GCP location was interpolated to a TIN and then resampled to a raster model at a 1 m resolution and then added to the 2014 surface (see Figure 7.8). The resulting surface is notably less smooth that the marginal correction surface, reflecting the strong dependence on the sparse distribution of residual measurements.

7.5. Results

In order to assess the approaches to treating the systematic errors described above, a DoDs were created for the 11 km test reach for each of the corrected models and compared to that obtained for the original, untreated DEMs. In this case, only changes that fall into the dry-dry (2015-2014) categories of change are quantified, in order to remove the unreliable wet channels from the analysis.

The DoDs were filtered using a 0.28 m minLOD, reflecting the replicate uncertainty analysis described in section 7.2. This threshold only accounts for changes that are

significant at overlapping 95% confidence intervals (i.e., 1 sigma uncertainty for bar surfaces = 0.1 m, therefore the propagated error at 2 sigma (2 x 0.1 m) determined as the sum in quadrature = $\sqrt{0.2^2 + 0.2^2}$ = 0.28 m). For comparison, changes determined for an unfiltered, untreated DoD are also shown. The tabulated results of the change detections are presented in Table 7.7.

Table 7.7 Tabulated metrics of change detection as produced by the GCD software (Wheaton et al., 2010) for the Delta reach DoDs.

	Raw Uncorrected	**Uncorrected (minLOD 0.28m)**	**Bank Corrected (minLOD 0.28m)**	**GCP Corrected (minLOD 0.28m)**
Total Area of Interest (m^2)	9,326,683	9,326,683	9,326,683	9,326,683
Total Area of Interest with Detectable Change (%)	100%	43%	47%	30%
Total Area of Erosion (m^2)	3,406,241	999,840	2,198,573	1,348,953
Total Area of Deposition (m^2)	5,920,442	2,964,442	2,053,434	1,414,669
Total Volume of Erosion (m^3)	791,485	503,552	1,157,616	663,141
Total Volume of Deposition (m^3)	2,144,270	1,752,960	1,036,203	671,775
Total Net Volume Difference (m^3)	1,352,785	1,249,408	-121,413	8,634
Average Depth of Erosion (m)	0.23	0.50	0.53	0.49
Average Depth of Deposition (m)	0.36	0.59	0.50	0.47
Percent Erosion (%)	27%	22%	53%	50%
Percent Deposition (%)	73%	78%	47 %	50%

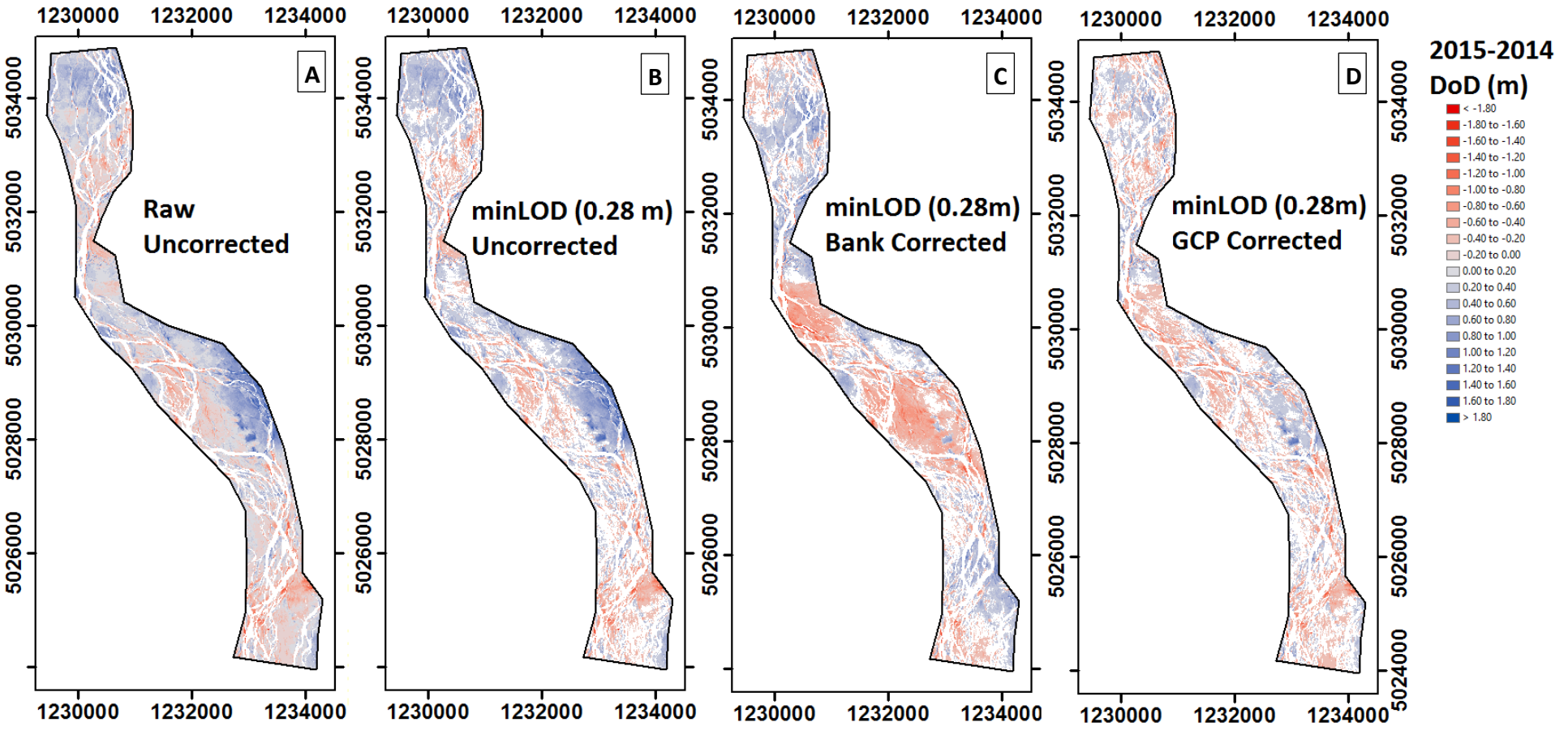

Figure 7.9 The raw and filtered DoDs produced during the process to mitigate systematic errors, where (A) is Raw uncorrected, (B) is minLOD 0.28 m uncorrected, (C) is minLOD 0.28 m bank-corrected and (D) is minLOD 0.28 m GCP corrected. The channel network has been removed for this analysis. *Projection: NZTM 2000*

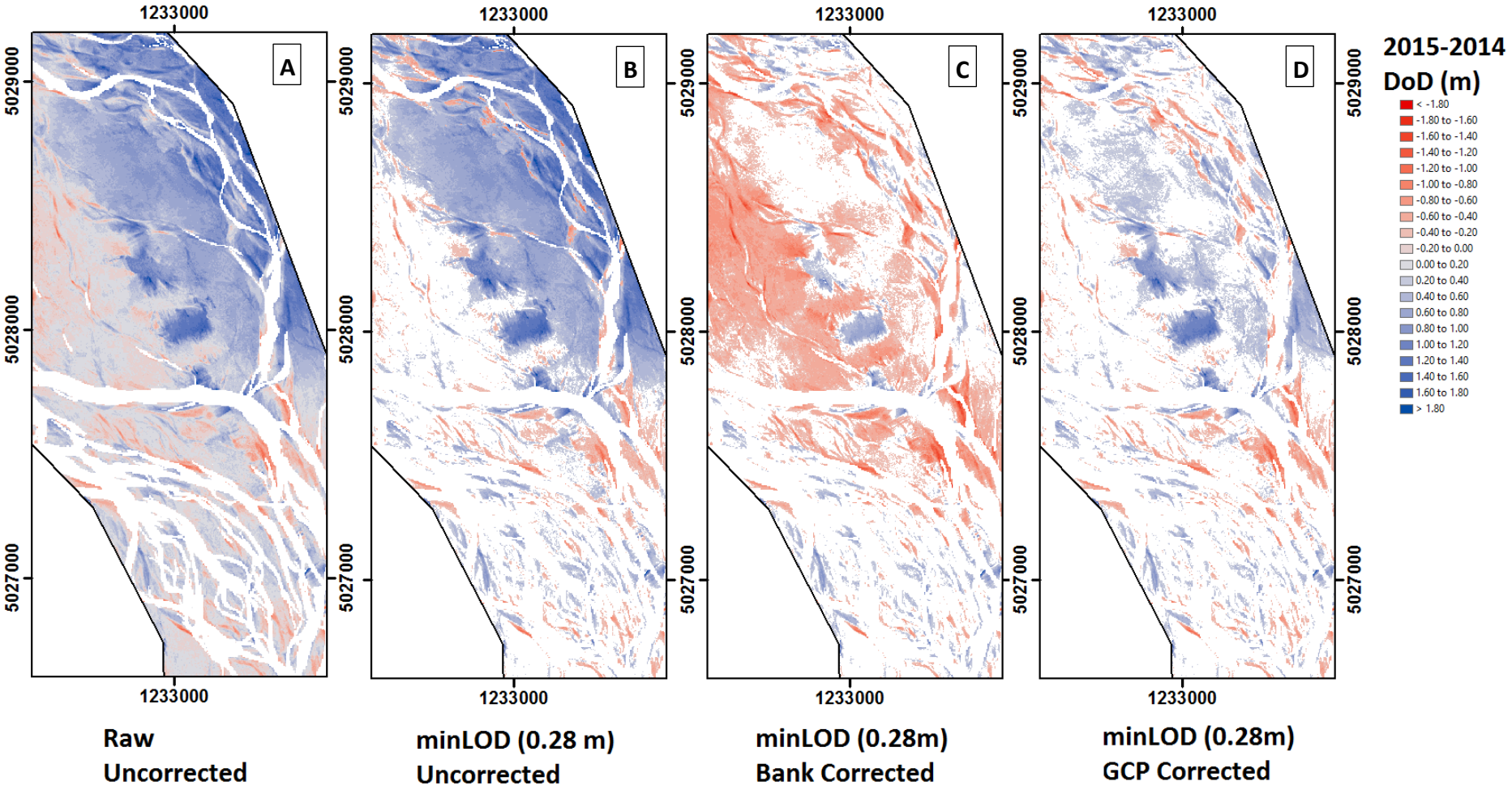

Figure 7.10 A close up section of the true left region that is prone to a miss match in ground control, thus exacerbating systematic error for the DoDs presented in Figure 7.9, where (A) is Raw uncorrected, (B) is minLOD 0.28 m uncorrected, (C) is minLOD 0.28 m bank-corrected and (D) is minLOD 0.28 m GCP corrected. *Projection: NZTM 2000*

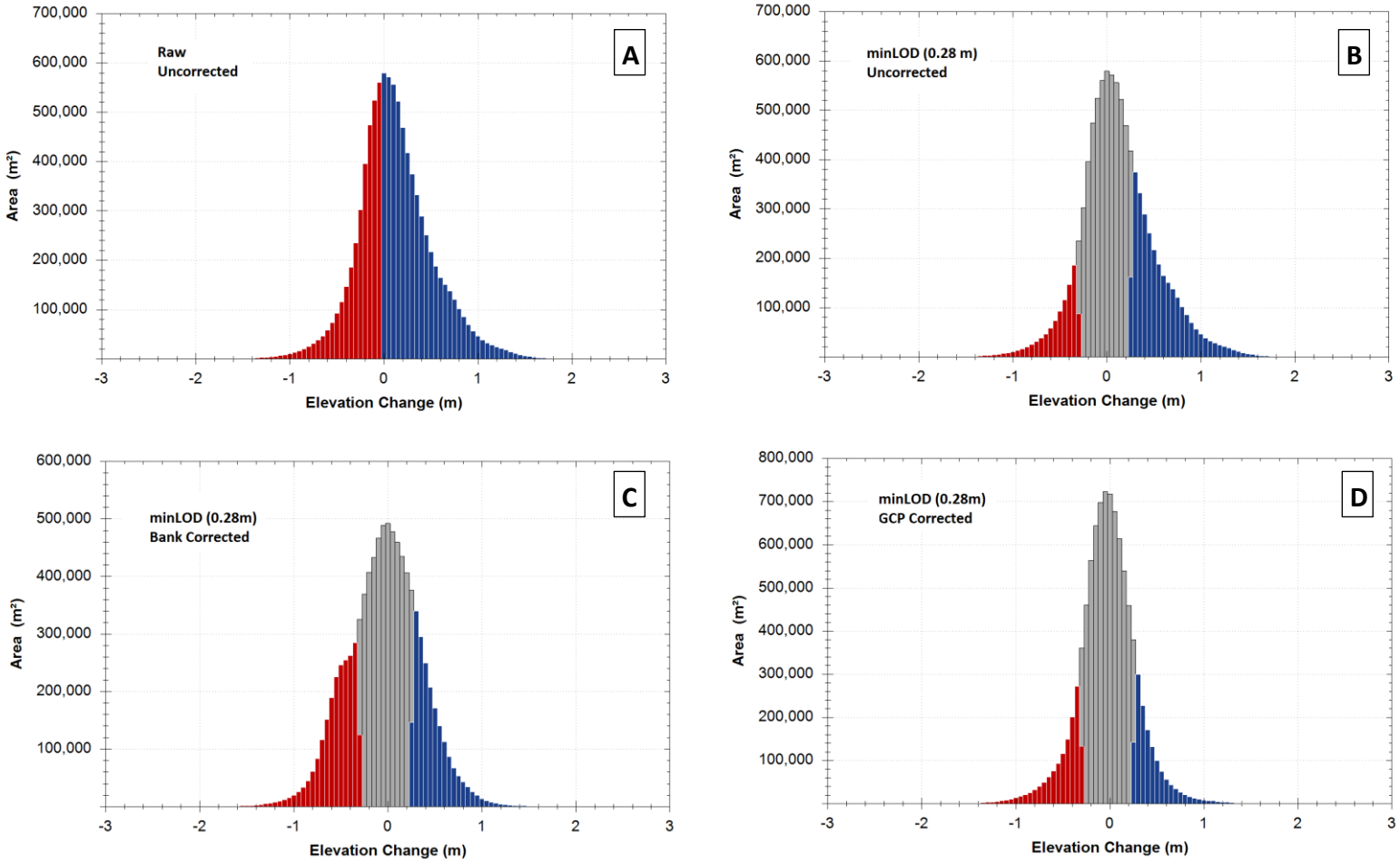

Figure 7.11 The aerial histograms for the change detection analysis of DoDs presented in Figure 7.9, where (A) is Raw uncorrected, (B) is minLOD 0.28 m uncorrected, (C) is minLOD 0.28 m bank-corrected and (D) is minLOD 0.28 m GCP corrected.

7.5.1. **Initial DoD**

While the test-case reach is, *a priori*, subject to a lower scale of systematic errors than the catchment wide DoD presented in Figure 7.3, the uncorrected model nonetheless exhibits significant bias. This is evident not just though visual inspection of the change map, but by examining the elevation change distribution shown in Figure 7.11. This reveals a strongly right-skewed distribution of change, which while centred about zero, has an imbalance of 22% erosion to 78% deposition. This reflects large increases in elevation on the true left, south of the Hillocks, which is associated with differing patterns of ground control between the epochs and a similar topographic rise at the upstream end of the model.

The zoomed in DoD provides clear evidence of the systematic error present, see Figure 7.10. The true right of the river is seen to experience equal positive and negative change, however this abruptly becomes dominated by surface lowering that masks the underlying geomorphological change. This coincides with a region or inconsistent ground control between surveys, resulting in the presence of bias in the peripheral regions.

7.5.2. **Bank Correction**

The bank correction analysis is prone to systematic errors, present within the DoD and the metrics of change presented in Table 7.7. The distribution of erosion and deposition is considerably more even, 53% and 47% respectively, however in this instance erosion dominates resulting in a net reduction in volume. The distribution of errors within the histogram (see Figure 7.11) remains unequal, however the positive tail has been reduced as seen from the initial model.

The spatial distribution of error, as seen in Figures 7.9 and 7.10 is once more unrepresentative of natural change. The immediate region downstream of the bridge at the Hillocks is represented as a surface lowering for almost 4 km in length. By zooming in we can see that at the very edge of the true left of the river, the bias associated to ground control inconsistencies has been removed. This, however, has been achieved by an unrepresentative surface lowering that has masked natural channel evolution for the majority of the active width.

7.5.3. **GCP Correction**

In contrast to the two previous DoDs the application of the GCP correction has resulted in an equal distribution of the aerial percentage of erosion and deposition. This is seen also within the histogram distribution, which is much more tightly constrained to a distribution about zero. The average depth of erosion and deposition has reduced from an initial 0.50 m and 0.59 m to 0.49 m and 0.47 m. whilst the net volumetric difference is reduced to 8,634 m^3. Furthermore the percentage of the area within the area of interest is reduced to 30%, a far more reasonable value when considering this analysis omits the channel network.

The spatial distribution in Figure 7.9 shows a reduced percentage of areas that relate to systematic surface elevation or lowering. In particular the true left region displayed more close up in Figure 7.10 is prone to less regions of supposedly > 1.8 m change. As a result it is possible to see the patterns of channel migration of those which were insufficiently inundated to be easily identifiable for removal within the orthophotos. Whilst the GCP based correction process employed is fraught with uncertainty, it has shown the ability to derive DoDs of realistic volumetric and depth change estimates of natural change between epochs

7.6. Discussion

7.6.1. **Internal software errors**

The replicate analysis described in section 7.2 is aimed to evaluate the latent uncertainties in the reconstruction of dense points, associated with the processing of the SfM bundle adjustment and SGM depth mapping algorithms employed in APS. These effects arise from a number of related processes.

First, the estimated camera geometry solved by the bundle adjustment is strongly dependent on the image information content (texture, structure). This affects the identification of keypoints (their number and reliability) and the subsequent correspondence analysis. In the experiments described here, by keeping the imagery constant but repeating the alignment process, the replicates revealed the effects of the random seeding processes that influence the keypoints generated and the images they

correlate. Variance in the results reflects the propagation of errors as images are incorporated and matched sequentially. The ultimate solution obtained thus depends on image order and the keypoints used. In this application, the bundle adjustment is solved using inner constraints (image information) only, although *a priori* information on the camera position is used to control the correspondence analysis by determining which photographs could theoretically be correlated due to their spatial coincidence. This information dramatically reduces the search complexity, simplifying an n^2 problem to one which is essentially linear, dramatically reducing runtimes. The benefits are not just computational however, as simplifying the parameter search enables more keypoints to be compared between relevant images, adding redundancy into the solution and enhancing the accuracy of the estimate. The approach is subtly different from a fully forward photogrammetric process however, where the camera pose is known and used to determine the camera centre from which 3D positions can be obtained intersection directly (with an appropriate interior camera model).

Once the camera orientation (interior and exterior) has been solved by the bundle adjustment, this geometry is used to undistort image pairs and match pixels along epipolar lines to generate the dense point cloud. This process is implemented using the efficient SGM algorithm, but as with keypoint identification and correlation, the quality of pixel matching will again be affected by the information content of the images. Thus in both processes, poor feature geometry and variability between images (i.e., exacerbated for dynamic features such as vegetation blowing in the wind, or moving water) will introduce noise into the resulting point cloud. Errors in the camera geometry will manifest in poor image alignment, which in turn will compound pixel correlation in the dense reconstruction.

These combined effects are clearly evident in the surface uncertainty calculations derived from the replicates. The complex, but well-textured and undeformable gravel bars exhibit comparatively low variability (SD = 0.1 m), whereas the dynamic water surfaces have considerable elevation variance (SD = 0.73 m). On closer examination, the number of tiepoints in the sparse cloud in these areas exhibit considerable variation, with gravel bars having an average of 0.8 pts m^{-2} versus water with 0.002 pts m^{-2}. It follows that the camera geometry over these areas is likely to differ in quality, and effect

which will degrade the subsequent depth mapping, which will also be compounded by the less structured image information.

It should be noted that the methodology of the 'reproducibility' of surfaces is only partial. Geomorphological change would normally be measured between at least two fully independent surveys. These would incorporate different camera networks, GCP placement and measurement (both in the field and lab) and potentially differences in sensor and allied instrumentation such as camera GPS. Any threshold for change detection from these results, therefore represents a conservative measure, and should be treated cautiously.

7.6.2 Bathymetry

No attempt was made here to apply a more sophisticated model to extract the submerged bed topography. Methods for deriving water depths (and with supporting information on water surface) and submerged bed levels do exist (see for example Dietrich, 2016), but are beyond the scope of this thesis. Additionally, it was recognized early in the research that conditions at the study site, in particular, the very high turbidity (with sediment concentrations exceeding 2 g/L) and surface waves, were unlikely to lead to successful bottom extraction using either optical reflectance (e.g., Williams et al., 2013) or through the water photogrammetry (Dietrich, 2016).

While the site conditions were far from ideal, it should be recognized that successful bottom extraction would also likely require differences in the camera network. Through the water photogrammetry has been shown to work best for low altitude (high resolution) imagery, collected with a polarizing filter and shot (ideally) with the sun behind the sensor (Dietrich, 2016). The results of the replicate analysis found that the wet channels had very high surface noise, resulting from spurious keypoint and pixel matching, which given the turbidity, reflected the water surface rather than bed topography.

For the purposes of later analysis, the major channels were therefore masked in the DoD and changes relating to only bar-bar transitions were considered. Unfortunately, this process obscures one of the key agents of geomorphological change within the system, i.e., bank erosion.

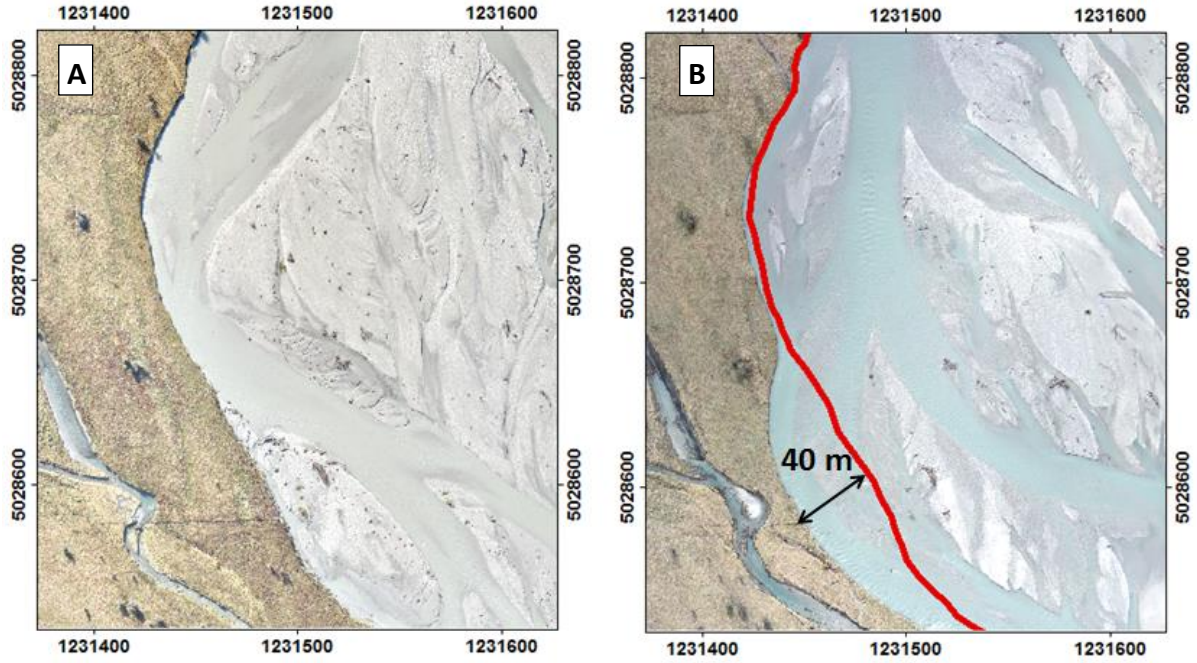

Figure 7.12. A) 2014 othophoto and B) 2015 orthophoto. Evidence of significant bank erosion experienced between surveys along the true right bank that will have been excluded from the analysis due to the difficulties in reconstructing inundated channels. *Projection: NZTM 2000*

Figure 7.12 provides high resolution, but isolated example of bank erosion occurring within the reach that illustrates over 40 m of lateral erosion between 2014 and 2015. Such high rates of bank erosion are evident along much of the lower true right of the Delta reach where, in places, the active fairway has migrated over 300 m westwards over the last five years. A comparison of the orthophotos from 2014 and 2015 suggests some 67,000 m^2 of floodplain was lost to the active channel within the year from just the lower 8 km of the reach. Assuming these losses are largely on the true left, this reflects an average bank retreat of over 16 m/yr/m. If we assume a typical floodplain top to channel bed elevation difference of ~1 m, this implies an additional 67,000 m^3 of unaccounted erosion, or another c. 10% of the estimated (corrected) erosion measured.

In addition to bank erosion, the wet-wet, wet-dry, dry-wet transitions in the active channel (e.g., Figure 7.13) are likely to represent a major source of geomorphological change.

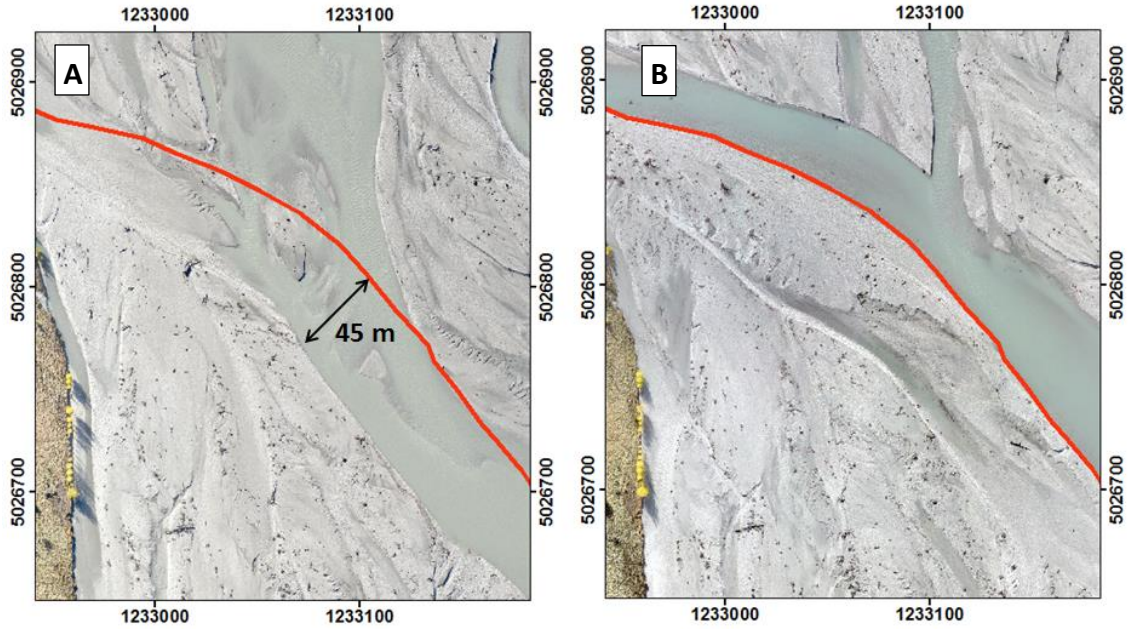

Figure 7.13 Evidence of within channel sedimentation processes between A) 2014 and B) 2015. *Projection: NZTM 2000*

Westway et al. (2003) analysed changes on the similarly sized Waimakariri River in NZ using traditional photogrammetry that incorporating bathymetric measurements. Segmenting changes into comparable classes (dry-dry, dry-wet, wet-dry and wet-wet), they observed that over an annual period (Feb 1999-Feb 2000), dry-dry changes comprised 64% of the volume of detectable erosion and 80% of the volume of detectable sedimentation. By inference, this implies a comparable loss of 36% and 20% of changes by neglecting the dry-wet, wet-dry and wet-wet classes. Following a major event in 2000, their analysis was repeated and found that dry-dry changes accounted for 58% of erosion and 64% of sedimentation by volume. Comparing these estimates to the Dart is difficult, but it is clear that the dry-dry bed transition dominates areal. Removing the wetted channels from both years, accounts for a total area of 3,422,892 m^2, reflecting 36% of the area of interest. However, it is likely that, in particular, dry-wet and wet-dry changes, which involve lateral migration or avulsion of anabranches will be associated with large changes in height and so may add disproportionally to the volumetric contribution. Any wider use of the results gained here should therefore be treated cautiously as they represent a significant underreporting of actual bed turnover and volumetric change.

7.6.3 Systematic Errors

The reconstructions derived in this research exhibit significant systematic errors that dominate the signature of change between the two surface models. These effects arise from the combination of inaccurate camera calibration in the bundle adjustment (both interior and exterior) and post-hoc optimization of the solution using a sparse network of ground control. The absence of either a theoretical model of these effects or an independent reference surface, effectively precludes precise quantification of the structure of these errors.

In order to facilitate an analysis of geomorphological change, two related procedures were developed that sought to correlate the models using areas presumed to be stable over time. This approach led to two spatially-explicit, correction models that were used to transform the 2014 survey to fit the structure of the 2015 survey. Two key caveats should be recognized from the outset. First, fitting the earlier survey to 2015 model, inevitably preserves the 2015 error structure within the two data models. This facilitates effective relative comparison, although does not rectify the error absolutely. Secondly, the derived correction assumes that the two models are horizontally aligned so that the net vector of error is purely vertical.

There is no way to independently validate the results of either correction procedure, however the analysis raises some interesting observations. First, by comparing the raw, unfiltered DoD with that subjected to the basic 0.28 m minLoD filter dramatically reduces the area of detectable change; down from 100% to just 42.5% of the study area (with wet class changes already removed). However, this simple analysis does not address the broad, systematic errors which persist in the filtered surface, with major elevation rises at the upstream of the reach and on the true right between 5,030,000-5,028,000 m N persisting. These features appear atypical of fluvial processes, although such activity cannot be entirely discounted. Application of the both correction models appears to remove these distortions, although a lack of consistency between the resulting patterns of change make it difficult to validate either model effectively. The bank correction model actually results in an increase in the area of detectable change, up to 46% and two new major areas of negative change appear in this model. The first of these is immediately downstream of the Hillocks (~5,030,500) and again centred at

approximately 5,208,500 N. Such broad patterns of negative change, are unlikely to be the result of erosion, which is normally associated with relatively narrow scour paths. Moreover, downstream of the Dart River Bridge at 5,030,500 the active braided fairway widens dramatically, which is likely to lead to regional reduction of shear stress and sedimentation rather than erosion. The GCP corrected model by contrast reduces the detectable change markedly, down to just 29.6% of the raw predicted changes. Again it is difficult to draw too many conclusions on the verisimilitude of the predicted pattern, although there are fewer areas of broad change irrespective of channel geometry. Interestingly, both correction models suggest that the reach is approximately in equilibrium topographically, with vertical rises in balance with falls. The average depth of changes is also similar, and close to a representative 0.5 m for both negative and positive changes, which is consistent of the bar top morphologies observed in the field.

From a retrospective perspective, one strategy that would have supported the correction modelling (indeed the whole reconstruction process) is to have used the same pattern of GCP placement in 2014 and 2015. As shown in Figure 7.14, there is considerable variability between the two surveys, with a deficit of control on the true left from 5,031,500 south to 5,028,000. Replicating the placement pattern would have provided a network of datum control that would be invaluable in ensuring effective correlation of the two models, and should be advised for future research.

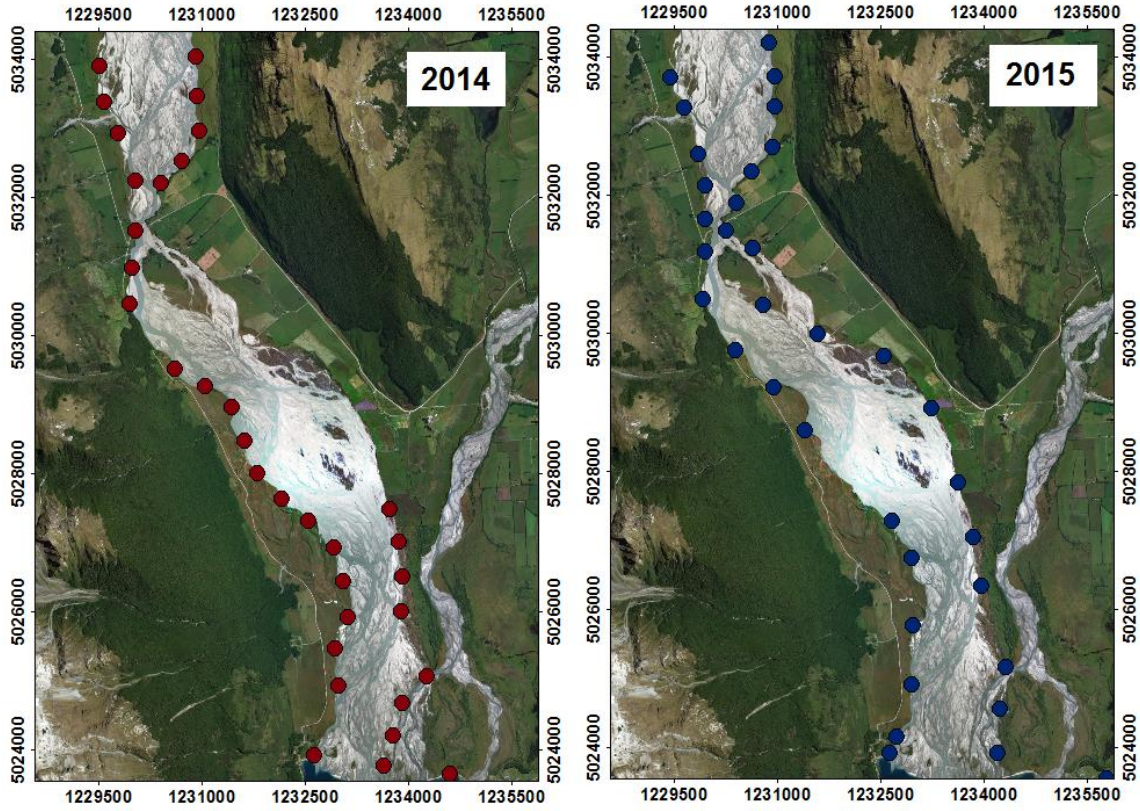

Figure 7.14 The differing extent of ground control in 2014 and 2015, highlighting the presence of gaps in 2014 which relate to the regions of largest systematic bias. *Projection: NZTM 2000*

7.7. Conclusions

The presence of severe systematic bias in the 2014 survey ultimately limits the interpretation of geomorphological change over the study period. Rather than quantify this uncertainty and incorporate it within a change detection framework, strategies were developed here to attempt to mitigate these anomalies by correlating the two surface models. This approach results in a significant refinement of the pattern of observed topographic change, although it remains impossible to externally validate these. While this result is disappointing, the key issues associated with the change detection arise from the low quality 2014 model, whereas the revised survey strategy employed in 2015 provides a good basis for ongoing research.

A conservative approach to quantifying local uncertainties in the reconstructed surfaces was implemented using replicate modelling of a test-study reach around the Hillocks. By repeating the reconstruction process with a consistent image and control networks, it was possible to see how variations in the structure of surface information captured by

the photography propagated into uncertainty in the vertical dimension. As expected, complex and deformable surfaces such as vegetation and water result are associated with significant variance by comparison with hard, well textured surfaces such as the gravel bars. The repeat framework provides an insight into the reproducibility of surface modelling, which was used to define thresholds for change detection. These suggest that for dry-dry comparisons, a level of detection of 0.14 m at 1 sigma is appropriate, comparing favourably with estimates found for traditional metric aerial photogrammetry (e.g., Brasington et al., 2003; Lane et al., 2003). While this uncertainty is above that obtained elsewhere for low altitude UAS SfM or terrestrial laser scanning (e.g., Brasington et al., 2012; Eltner et al., 2017), this scale of measurement is sufficient to quantify fluvial dynamics effectively and can be achieved at lower costs than competitor survey methods. It should, however, be recognized that the LoD determined here does not incorporate the full source of uncertainties that effect reproducibility between fully independent surveys. Further studies at the full scale of the experiment here, encompassing alternate camera and control networks and even different sensors and positioning systems is required to evaluate this.

8. Chapter 8: Discussion

8.1. Introduction

Chapters 4 to 7 have shown the ability of SfM to derive landscape scale (10-10^2 km^2) terrain products for the purpose of geomorphological change detection analysis. This has been achieved through a refinement of an initial SfM survey in 2014 that employed a survey strategy in-line with standard UAS procedures documented within the literature. Subsequent analysis of the pipeline process within a scaled setting led to the development of an updated data acquisition and processing strategy, orientated around the use of a dual camera convergent image network with *a priori* pose estimates. Subsequent modelling of newly acquired survey data in 2015 using this approach yielded significant dividends in terms of model performance. Mean Absolute Errors based on check points were reduced from 0.44 m in 2014 to 0.23 m in 2015 with a noticeable reduction in model bias introduced by post-hoc optimization of the bundle adjustment. Despite these improvements, a reach-wide analysis of geomorphological change remains problematic given the inherent systematic errors within the 2014 survey. To combat this, a spatially variable mitigation strategy has been proposed to tie the surveys together enabling the assessment of a 17 km^2 area in the lower reach of the Dart River.

Specific discussion relating to Chapters 4-7 are contained within each section and so the focus will fall here on the broader findings of the research. The over-arching aim of this thesis has been to assess the ability of SfM to derive landscape scale terrain products which have sufficient reproducibility to enable the analysis of change detection. This discussion will focuses on the following themes:

- The use of convergent camera networks and *a priori* pose estimation to enhance the reliability of the SfM bundle adjustment;
- A brief overview of hypothesized future improvements to the catchment scale SfM pipeline, which for practical reasons have not been explored further within this study;

A final point to consider when contextualizing the outcomes of this thesis is the inherent differences between typical small scale UAS implementations of SfM and the landscape scale approach proposed here. The survey refinements implemented between 2014 and 2015 result in a significant reduction in systematic error as they are designed primarily to mitigate for sub-optimal exterior constraints. Whilst the doming effect is prevalent

within nadir orientated UAS SfM surveys, the impact of this distortion is felt most severely beyond the GCP perimeter (James and Robson, 2012). As such, typical survey strategies constrain the study region to this theoretical perimeter enabling the focus of a majority of work to fall upon survey strategies relating to local refinements (Turner et al., 2012; Micheletti et al., 2015; Clapuyt et al., 2016). In contrast, the emphasis of this thesis has been toward minimizing the global systematic issues that arise in the absence of adequate ground control at the expense of refining fidelity at the local scale.

8.1. Convergent camera networks

Developing 'landscape scale' or 'wide-area' models encompassing 10^1-10^2 km^2 using SfM Photogrammetry remains challenging. Difficulties range from the shear logistical task of acquiring aerial and ground data over large areas, to complex environmental conditions such as inundated areas and mountain topography, to the computational resources required to solve the scene structure and dense reconstruction with a large image network. These themes were examined in detail in Chapters 4 and 6 (see Table 4.1 for a summary). A key conclusion from this research is that errors in the calculation of the 3D structure cannot be readily corrected through updating the block using a sparse network of GCPs. The incorporation of GCPs as exterior constraints in the optimization of the bundle adjustment serves to lock the solution accurately at these locations, but at the expense of biasing the interior and exterior camera calibration and creating systematic distortions of the resulting 3D structure that are difficult to unravel. For smaller studies, such as UAS surveys covering areas 0.1- 5 km^2, it is often possible to circumvent this problem (in fact neglect it entirely) by deploying a dense, regular control network. However, as the scale of the study increases, so inevitably does the achievable density of ground control and the potential to mitigate the misestimation of the camera geometry by retrospective correction.

An important aim of the work described in Chapters 5 and 6 therefore was to investigate data acquisition and processing strategies to support robust initial estimation of the camera geometry, thereby relaxing the need for subsequent correction with dense exterior constraints. A number of recent studies have highlighted the potential to use convergent camera networks to reduce systematic errors originating from

compensation in the bundle adjustment (James and Robson 2014; Woodget et al., 2014; and Dietrich, 2016). These build on the empirical analysis of Wackrow and Chandler (2011) who found that stereopsis based on oblique imagery resulted in improved estimation of depth and allowed accurate recovery of known interior camera geometry. This study represents the first large scale evaluation of this approach, here using a nadir 30^{o} aft-nadir camera set-up to obtained a combination of vertical and oblique photography and was found to: 1) improve interior calibration estimations; 2) improve pose estimation; and 3) reduce the reliance on ground control.

8.1.1. Improved interior calibration estimation

Systematic distortion has previously been linked to the presence of errors in the estimation of the radial distortion parameters, principally the first significant K1 parameter (James and Robson, 2014). This distortion is manifest in a doming pattern of block distortion (James and Robson, 2012) and was clearly observed here in the reconstruction of the debris fan in 2014 (Chapter 4). This structure of distortion reflects the extent and distribution of ground control. However, for the corridor mapping strategy applied to survey the entire reach, the pattern of distortion observed is far more complex characterized by low frequency wave-like structures. Whereas the simple dome effects have been linked to misestimation of the K1 parameter, the complexity of the pattern observed here is hypothesized to result from the simultaneous misestimation of the K1, focal length and 3D camera pose. To facilitate the ease of explanation, the following shall focus primarily on the impact of radial distortion and focal length, before latterly addressing pose estimation.

Results in Chapter 5 examined the effect of altering the focal length calibration under a range of different convergent camera networks. The use of a fixed calibration with a deliberately erroneous focal length provided the means to assess the structural effect of incorrectly estimating this parameter. The resulting 3D structure exhibited a bowl pattern in contrast to the doming observed in similar experiments examining the sensitivity of radial distortion, K1, parameter. The degree of 'bowling' by the Focal length is found to be less than the 'doming' effect unless significant alteration to the calibration is applied, e.g. >5%. In the case of the debris fan a 0.67% degree of change is experienced

in the focal length during optimization, as opposed to 0.98% in K1. This ultimately results in a doming pattern during the optimization which is propagated into the final model.

By comparing the change in finalized radial distortion and focal length parameters in the 2014 and 2015 surveys, it was possible to derive estimates of the reliability of calibration. Under the assumption that convergent imagery provides a more reliable calibration estimation, the degree of change in calibration parameters during optimization is a proxy for the magnitude of error in the initial calibration. Between the 2014 and 2015 surveys there is a parameter change of 5.19% (radial distortion) and 2.6% (focal length) respectively, which compares favourably to the 5.75% and 2.02% recorded in Chapter 5 between nadir and convergent flume reconstructions. The convergent camera network therefore was found to provide more significant improvement in radial distortion estimation than that of focal length. This, however, fails to explain the presence of the wave form pattern in the optimization phase.

The changes in interior geometry parameters during the optimization provide an insight into the dependence upon the ground control. Results from the 2014 survey, however, indicate a surprisingly low sensitivity of the radial distortion parameter during optimization in contrast to that observed in 2015 (1.58×10^{-2} % vs. 0.15%). The extremely low rate of correction is potentially symptomatic of wider inaccuracies in the global solution. This is not reflected in the focal length parameter (1.79×10^{-2} % vs. 1.04×10^{-2} %) for 2014 and 2015 respectively, however, the interaction between parameters during the bundle adjustment reduces the reliability of these estimates.

Analysing the optimization of the reach-wide surveys highlights two key changes between the years. These are a reduction in the magnitude of error and an alteration in the shape of wave, perhaps best expressed as a reduction in wave height and increase in wavelength, see Figures 8.1 and 8.2. The following section (see 8.1.2) addresses the magnitude of change which appears linked closely to pose estimates. This is best observed in the dramatic shift in wave height and wave length between optimizations for Scan C2 and Scan D (Figure 8.1)

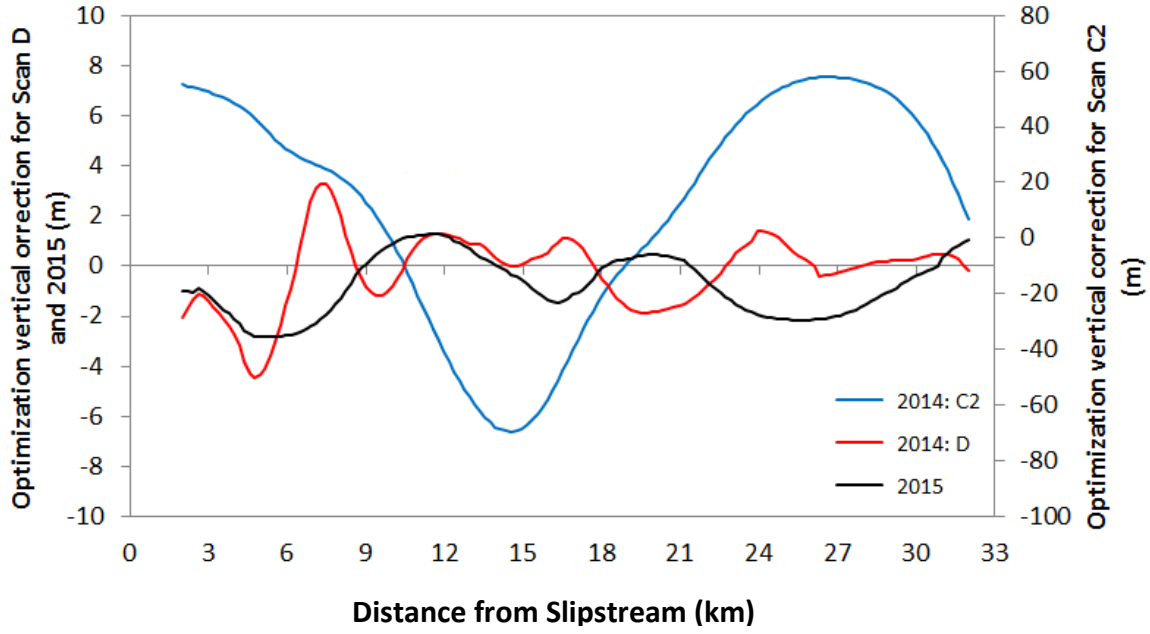

Figure 8.1 Representation of longitudinal wave pattern of optimization seen in the Dart survey, corresponding to Figure 8.2. The amplitude and effective wavelengths are varied for each reconstruction, most noticeably a reduction in amplitude and increase in wavelength between 2014 and 2015.

The change observed between Scan D and the 2015 survey is proposed to be a result of both further improvement in pose estimations and the impact of convergent imagery.

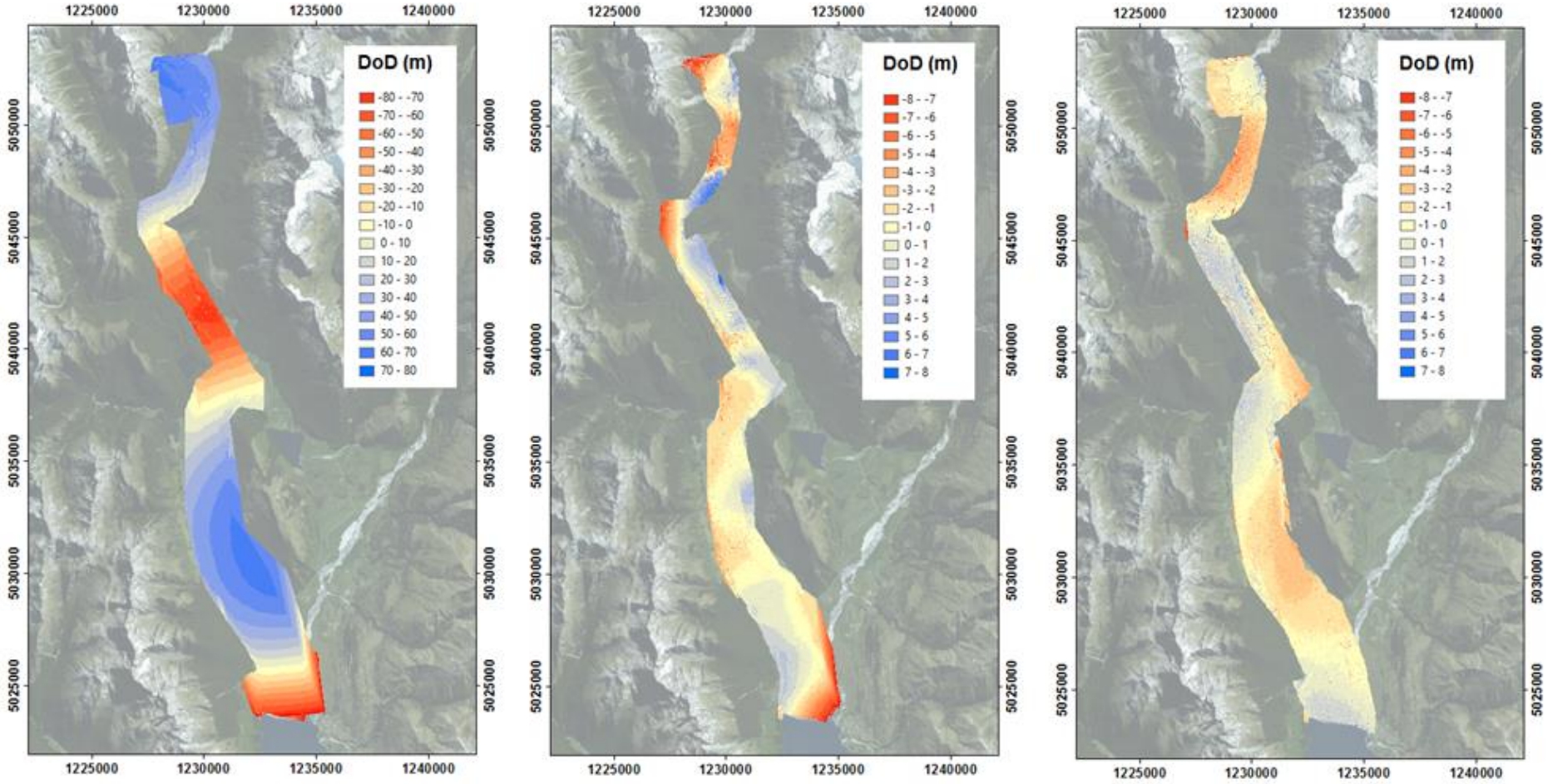

Figure 8.2 The optimization phase within APS for SCAN C2 (2014 initial), SCAN D (2014 refined) and the 2015 survey. The initial scan C2 is dominated by poor pose estimates that result in a surface 'wave height' of 80 m with a mean wavelength of approximately 15 km. In comparison improved pose estimates reduce the 'wave height' to 8 m with in SCAN D with a mean wavelength of approximately 5 km. The 2015 survey sees the 'wave height' reduced further to just 4 m with a mean wavelength of approximately 11 km. It is proposed that the degree of change experienced in this optimization is analogous to a proxy for model accuracy. Note the difference in scale between years. *Projection: NZTM 2000*

An insight into the origins of this waveform is provided by comparing the 2014 local model of the debris fan to the catchment wide survey in 2014. As noted previously, the debris fan model exhibits a dome-type deformation when constructed as a regional block, although such patterns are not present in reach-scale model (see section 4.5.1.). The two major differences between these reconstructions are the quality of pose estimates and distribution of ground control. For the local model, pose estimates are derived via a regional block of ~500 cameras with seven GCPs. In contrast, the reach-wide block comprises of ~4000 estimated from 81 GCP's. However, this latter model is seeded initially with poor quality exterior constraints derived from Google Earth (see heuristic workflow in Chapter 4). The distribution of ground control for the local model is, meanwhile, accurately measured, comparatively dense and evenly distributed across the debris fan. However, regardless of the distribution of control, it is clear that: a) the wave pattern is present in the initial sparse cloud before the addition of ground control and; b) that the change in wavelength between 2014 and 2015 is not simply linked to the density and placement of ground control. It is proposed, through process of elimination therefore, that the wave pattern is a result of accumulated errors that arise in the bundle adjustment as images are incorporated sequentially and inconsistent pose and interior geometrical estimates are iteratively updated into a sub-optimal solution. By contrast, the inclusion of oblique perspectives in the 2015 block provides a stronger geometrical framework for the simultaneous calibration of the interior and exterior geometry and therefore, a more robust solution of the bundle adjustment (Figure 8.2).

8.1.2. Improved Pose estimation

The significant advantages of correct pose estimation with regards to positional data in the *x*, *y* and *z* are discussed latterly in relation to camera GPS. This section instead focuses on the compensation between the internal and external camera parameters for nadir and convergent camera networks. The controlled nature of the flume in experiments enabled a detailed analysis of pose estimates for nadir-only and convergent image networks. Additionally, the 2015 field survey also enables the direct comparison of camera networks that obtain imagery from the same physical space but with different combinations of viewing geometries (i.e., nadir with or without oblique views). This can be seen clearly from the nadir and convergent reconstructions generated of a bar

complex in proximity to Chinaman's Bluff in 2015 (section 6.2.2.). The reconstruction consisted of 165 nadir images as a nadir only network and as part of a 350 image convergent block. These were georeferenced by four GCPs and optimized to provide the final estimated camera positions. Comparing the 'identical' nadir images across different reconstructions displays an offset in image location between identical cameras in nadir only (A) and convergent networks (A'), see Figure 8.3.

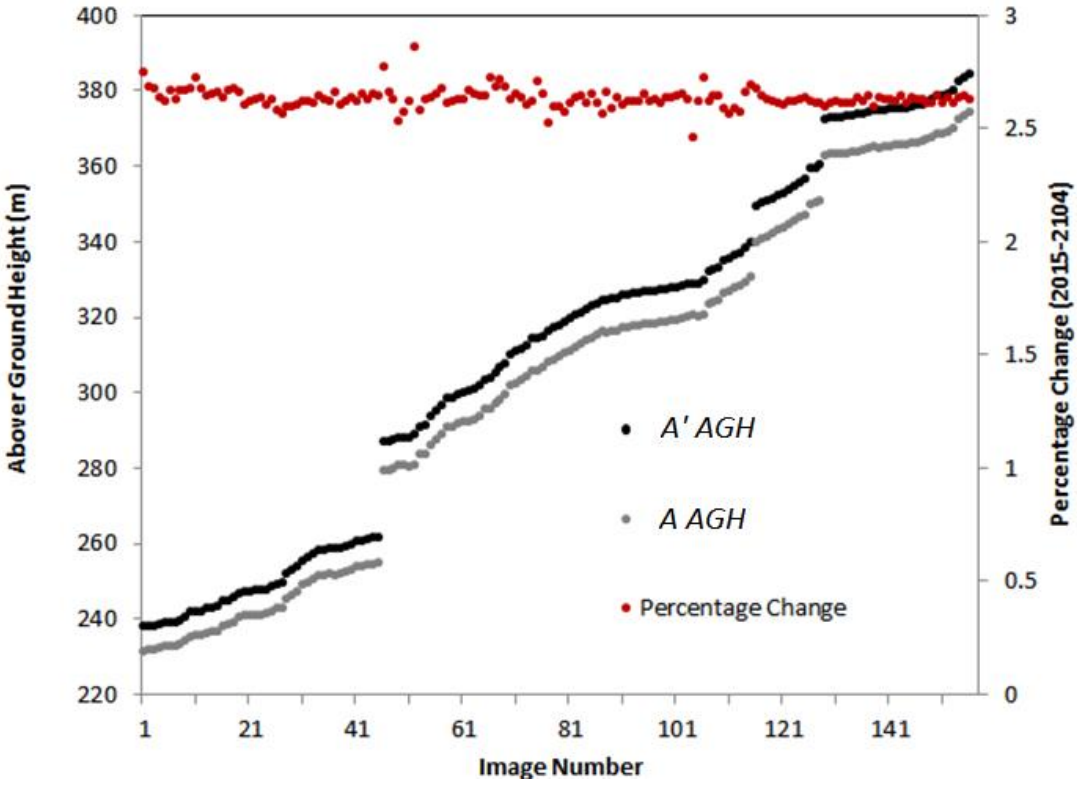

Figure 8.3 The difference in height between nadir imagery when part of a nadir-only camera network and as part of a convergent camera network.

Within the flume the vertical offset *d* was found to average 0.03 m across 32 cameras, corresponding to 2.5% difference in height between the two networks (convergent higher). Using the simple geometrical scaling relation between the laboratory and field (see section 5.2.1) this offset should scale to a difference of 7.8 m, which compares well the difference of 8.1 m obtained by repeating this experiment with the field scale data (Figure 8.3). This variation in camera position is principally limited to the vertical dimension, with the much lower differences obtained for the horizontal dimensions, of just 0.15 m and 0.08 m respectively. These results therefore corroborate Wackrow and

Chandler's (2008; 2011) analysis of depth estimation for convergent stereo pairs, and suggest that similar benefits apply here with multi-view stereo reconstruction.

Finally, these differences provide an explanation of pattern of parameter compensation observed. As the nadir-only camera network is estimated to be (on average) some 8.1 m (or 2.5%) lower than the convergent network, optimizing the bundle adjustment to fit the observed exterior constraints (i.e., GCPs) results in a different set of interior camera estimates. This is achieved, principally, through adjustment of the focal length and to a lesser degree the K1 radial distortion parameter.

8.1.3. Reliance upon ground control

The improvements in the SfM bundle adjustment reduce the reliance upon ground control to locally correct the interior and exterior orientation. In this research, given the lack of a high resolution reference topographic dataset, the difference between the pre- and post-optimized models has been used to understand the quality of the initial 3D reconstruction. However, while optimization improves the solution obtained by the bundle adjustment locally, refining the camera orientation to fit the exterior constraints (GCPs) introduces new systematic errors that propagate between the GCPs resulting in complex distortions of the modelled surface. An insight into these systematic distortions is also captured by the pre- and post-optimization DoDs as illustrated in Figure 8.2.

As SfM involves the simultaneous solution of interior and exterior camera orientation, errors in the resulting bundle adjustment are likely to be commonplace due to the complex nature of the parameter space search and the limited constraints. However, as most applications are relatively small scale (areas < 1 km^2) these errors are readily obscured through the use of dense networks of exterior constraints during optimization. Up scaling the methods to larger problems and remote, hostile environments as the Dart, comes at the expense of a much reduced network of control so that post-hoc optimization of the model is not only less effective, but as seen here, can further complicates the structure of errors. The difference between solving the SfM problem locally and regionally is effectively illustrated by comparing the reconstruction of Glenorchy, modelling from the local block in Chapter 5 and the reach-scale model in

Chapter 6. Table 8.1 shows the maximum observed ICP errors from these two models, and shown here for both nadir-only and convergent camera networks (Table 8.1).

Table 8.1 Max E values for both the control experiment at Glenorchy and across the entire reach. It is important to recognize the reach-wide values are undertaken under different survey designs and as such the degree of error reduction may not solely be a result of the convergent camera network.

Study	Nadir MaxE (m)	Convergent MaxE (m)
Glenorchy	0.40	0.18
Reach-wide*	2.67	0.63

In this case, the nadir photography model can be effectively corrected given the density of ground control, with the maximum ICP errors of only 0.4 m. However, the much sparser network of constraints in the regional model is results in significant distortion across this peripheral area of interest, giving rise to significant errors of over 2 m. The use of convergent imaging networks in 2015 was aimed to increase the accuracy of the initial bundle adjustment, and so reduce the reliance on post-hoc optimization of the model. Encouragingly, this approach resulted in maximum ICP errors (0.63 m) in the regional model that are broadly comparable (0.4 m) with those in the optimized local model of Glenorchy.

These results indicate that convergent camera networks provide a useful approach to enhance the information used to solve the bundle adjustment, effectively rebalancing the inverse solution. There are alternative approaches to achieve this, and there is considerable recent interest in the potential for direct georeferencing with SfM. This involves the use of high quality camera pose information that is used to determine camera centres and orientation *a priori* (see Smith et al., 2016 for further discussion). While this approach may offer some potential, it is clear that direct georeferencing is likely to require very high precision measurements of position and attitude. This may be possible for local networks, close (1-5 km) to a reference GNSS receiver, however as the problem is again up scaled, the quality of differentially processed GNSS positions falls and inevitable loss of radio lock will require the use of PPK position solutions that incorporate the risk of solution failures should the remote sensor lose lock. The multi-sensor camera system used here, provides a comparatively risk free and lower cost

solution, generating convergent photography that supports accurate solutions using inner constraints only.

8.2. Camera GPS

A major advantage of SfM over traditional photogrammetry is the ability to use non-metric low cost un-calibrated cameras (Westoby et al., 2012; Fonstad et al., 2013). The introduction of *a priori* camera pose information to help constrain the bundle adjustment does, however, offer significant advantages to landscape scale studies. Unlike the direct georeferencing approach described above, in this study a code-only, commercial grade GNSS receiver was used that logged latitude, longitude and ellipsoidal height only to the EXIF image file. The potential benefits from logging GPS data during the image capture are broadly two-fold: (i) increase accuracy in the bundle adjustment by constraining the search for conjugate images and therefore enhancing the accuracy of a correct local minima and; (ii) reduce the computational demand, and so enable the management of larger image sets; the use of refined processing settings; and increased image quality. This thesis has demonstrated that by collecting camera GPS data in this way, it has been possible to model a large area comprising 35 km^2 as a single photogrammetric block, obfuscating the need to match smaller, overlapping blocks covering the same area.

8.2.1. Improved Pose Estimation

Section 8.1 discusses how improved estimates of pose can be support reduction of systematic errors. Here, the impact of switching from a heuristic approach based on iteratively updating camera positions used in 2014 (Chapter 4) compares to their direct measurement using GPS in 2015 (Chapter 6). In the 2014 survey, a set of low quality exterior constraints were used to seed limited area models (the Gorge reach) that were iteratively refined to generate improved estimates of camera positions to support a revised alignment (bundle adjustment). By contrast, the direct GPS measurements used in 2015 provide independent, although certainly not error free estimates of position. Indeed, the positional drift associated with high survey flight speeds necessitated a time consuming manual correction of positional information. Nonetheless, the resulting

difference in the degree of camera positional error (a priori versus posterior position) is shown in Figure 8.3.

Figure 8.4 compares camera errors for the initial 2014 model (Scan C) with Scan D, in which the camera positions from Scan C were used to re-estimate the bundle adjustment. Both of these 2014 models are then compared to the camera errors associated with the convergent image model derived in 2015. Much like the optimization DoDs shown in Figure 8.1 these models of camera are a useful proxy for the reliability of the initial and final reconstruction. In turn, the increasing accuracy of pose estimates from 2014 to 2015 results in a significant reduction in camera change during the optimization. Further studies with the use of more reliable GPS systems as suggested by Smith et al. (2016) and covered in section 8.1.3 would be advantageous to further improve the accuracy of results. However, even without such data, it is noteworthy to compare just how effective the iterative approach to updating camera positions and resolving the bundle adjustment can be. The final average difference in camera position in Scan D is just 2.14 m, whereas Scan C exhibited numerous cameras with differences of >80 m. Whilst this is impressive, it indicates a significant amount of compensation is occurring between the internal and external parameters during the final stage of the bundle adjustment and hence increasing the reliance on ground control.

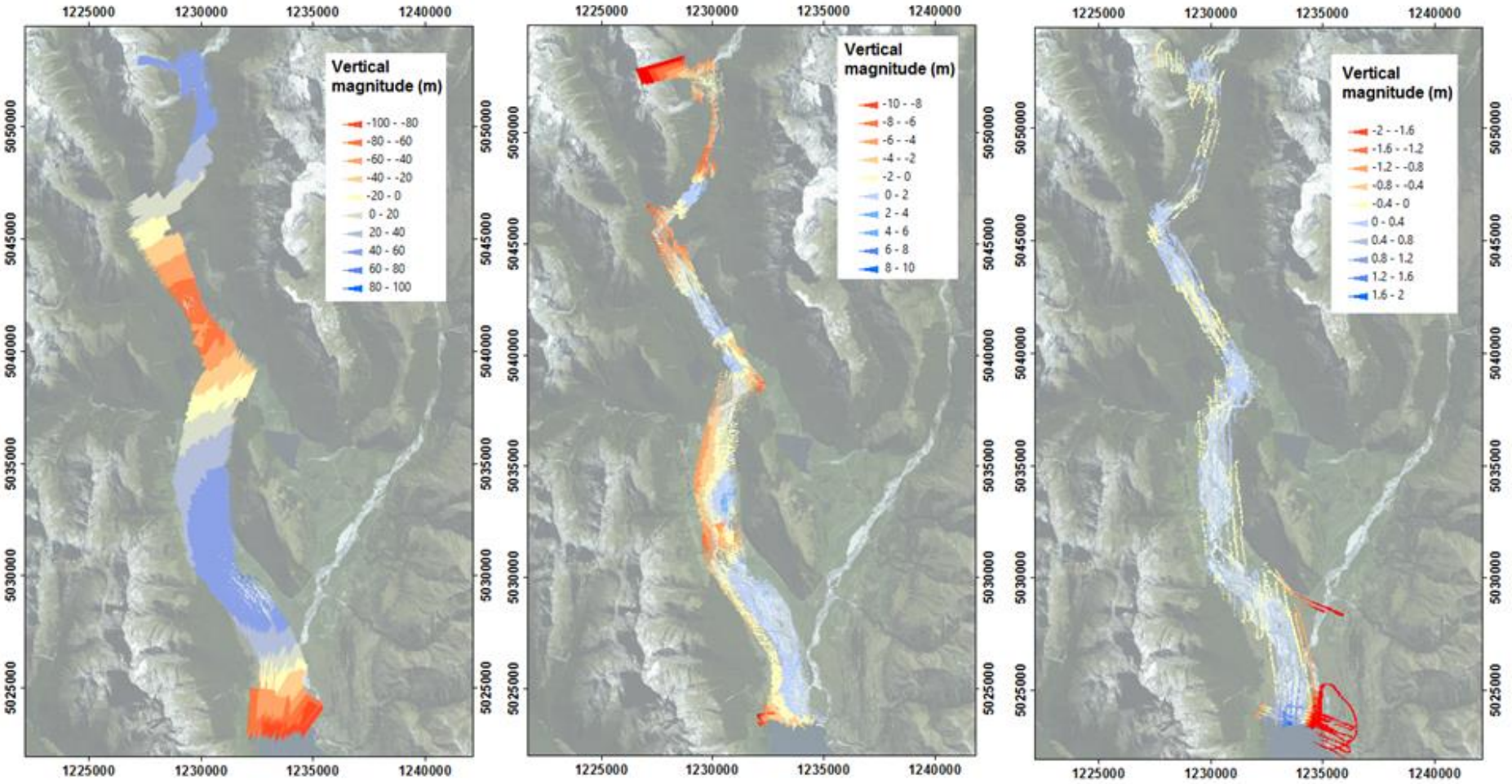

Figure 8.4 The magnitude, vertical direction and horizontal direction of camera change during the optimization process for SCAN C2 (2014), SCAN D (2014) and the 2015 survey. The colour ramp is indicative of the vertical direction, whilst the size of arrow is the total magnitude of change and the arrow direction is the *x y* change. Note the difference in scale between years. *Projection: NZTM 2000*

8.2.2. Reducing the Computational Overhead

The most practical reason for using a camera GPS is the potential to reduce the computational burden of data processing. The computation demands of SfM have been noted in a number of previous studies (Harwin and Lucieer, 2012; Smith et al., 2016) and in the past, limited its application to model regional landscapes (Dietrich, 2016). These computational constraints have also precluded a comprehensive analysis of the settings used in the global solution to the bundle adjustment.

Operationally, it remains important to capture imagery with at least ~60% overlap (Fonstad et al., 2013), and regions of complex geometry will benefit from even greater frequencies to improve redundancy (James and Robson, 2012). This is especially important for large-scale surveys where costs of chartering helicopters necessitate a fully redundant approach to data acquisition given the costs of further surveying. Along with sufficient image numbers, the quality of image sensor is vital to ensure the identification of keypoints (Furukawa and Ponce, 2010). The image quality – sharpness, resolution and exposure – is particularly important to maximize the image information and mitigate the impact of high flight speeds and heights (Micheletti et al., 2015). The initial quality of images is doubly important if reduced resolution settings are used (APS medium or low accuracy settings during alignment) which result in a reduction in keypoint quality.

Negotiating these survey concerns comes at a computational cost through increased imagery and increased image file sizes. These combine to increase potential keypoints within an image that must then be assessed for correspondences across increasing image numbers, which themselves contain an increased numbers of keypoints. To manage this increasing demand, *a priori* camera positional data allows the use of the pre-selection of imaging pairs under the Reference alignment setting in APS (see section 3.4). Computationally this results in only the cross comparison of local images as opposed to the entire imagery set. The result this has on processing times may be seen in Table 8.2.

Table 8.2 Alignment times within APS for comparative scans. The 2015 survey is omitted as the Disabled setting was unable to reconstruct. The reach-wide 2014 survey is misleading as the disabled reconstruction was only achievable using the Low alignment settings, whereas every other scan was performed under the high alignment setting. All reconstructions were carried out using a Dell Optiplex Dual Core I7 with 16 GB RAM and a 1 GB NVIDIA graphics card.

APS alignment setting	Disabled	Reference
Debris Fan 2014	5 hours 24 minutes	35 minutes
Reach-wide 2014	28 hours*	7 hours 42 minutes

These two examples illustrate the significant reduction in processing time for the bundle adjustment. However, while these are marked improvements, perhaps more noteworthy, is that due to the sheer size of the 2015 image set, it was impossible to reconstruct the 3D structure under the disabled alignment settings.

8.3. Opportunities for further survey refinement

The results from this thesis have identified a number of areas for potential future improvements when considering landscape scale SfM photogrammetry. These include:

8.3.1. Consistent GCP placement

This study has focused on minimizing the reliance on ground control, a consequence of which is a lack of attention to the role that this may play in shaping landscape scale SfM studies. Considerable attention has previously been paid to the effect of varying GCP density (Turner et al. 2012), however, only Clapuyt et al. (2016) has considered the role of GCP distribution on the repeatability of reconstructions. Unfortunately, this study is undertaken over a small area, and over relatively simple study geometry.

Using consistent GCP distributions offers a number of key advantages which are only apparent when considering repeat surveys. First, it offers a consistent framework for model optimization. While this will not necessary mitigate structural errors, it should result in their reproduction, thereby enabling differences between models to be quantified more effectively. Second, similar GCP networks also provide a comparable framework for model validation, allowing key differences in the performance of reconstructions to be identified. Third, there is scope to consider the use of fixed GCPs

(or stable nature features) which could minimize survey times and enable routing modelling of areas.

8.3.2. Applying a global image mask

Chapter 6 highlighted the potential to use a mask to remove the peripheral regions of individual imagery to reduce noise in the sparse cloud. The benefits of such an approach are twofold: i) an increase in the accuracy of tie points, as regions at the edge of the FOV are eliminated; and ii) a reduction in the image area/keypoints facilitates the use of additional imagery or reduced computational times.

Theoretically, this could lead to an improvement of tie points comprising the sparse cloud, with reduced reprojection errors and a resulting improvement in the camera calibration. The relative scale of this improvement compared to adopting convergent camera networks and/or camera GPS is likely to be minimal, however it is easily implemented and worthy of further consideration. In APS, images can be masked in user-defined batches, allowing flexibility in the degree of rigor preferred. This mask may be iteratively increased until a threshold is reached at which point reconstructions fail.

8.3.3. Iterative camera pre-calibration

Results from Chapter 4 described the use of iterative camera calibration to minimize errors at peripheral model locations. This is particularly attractive as it appears to further reduce the reliance on ground control. Unfortunately, the strategy used to derive the seeding calibration is limited, as ultimately the process acts to propagate uncertainties, building errors in the process. The quality of the initial estimate is also dependent upon the distribution of ground control.

Future work may prove more productive as improved pose estimates become available with the advent of RTK camera positional data. Furthermore the continual improvements made to subsequent editions of APS may provide more robust to the seeding of the bundle adjustment with poor exterior constraints. Beyond APS a growing range of software exist that require benchmarking and further analysis, and may offer improvements (see Table 2.1, Section 2.5).

8.4. Summary

The survey refinements adopted through this thesis have resulted in a means to capture landscape scale SfM terrain models. This has been achieved primarily through improving the initial sparse cloud geometry within APS and thereby minimizing the reliance upon post-hoc optimization of the model using ground control in a non-linear correction of camera geometry. There exist a number of ways in which to improve further surveys, however, the focus must be on improving reproducibility of results as opposed to accuracy.

9. Chapter 9: Conclusion

9.1. Overview

The SfM pipeline has demonstrated considerable potential for the future terrain surveying of large fluvial systems. Chapter 6 represents the first application of the SfM pipeline at such scale, conducted as one photogrammetric block with the use of a convergent camera network. The work undertaken here has mitigated the host of logistical challenges faced in wider scale SfM studies which result in both increased random and systematic errors. Whilst braided river morphology provides a host of unique challenges relating to the highly mobile and labile banks, significant wetted area and regions of high contrast, a number of issues prevail that would apply to all landscape scale surveys. Most significantly of these are the practical constraints placed on ground control distribution, the high flight height and speeds required to capture these areas and the resultant computational burden of the significant imagery numbers. This thesis has addressed these issues to both reconstruct the topography of the Dart River and perform a rudimentary geomorphological change detection analysis. The outcomes of this thesis may be judged against the three primary objectives identified in Chapter 1:

Objective A) To develop a field-to-data product workflow for 3D image-based topographic reconstruction of large fluvial systems.

Despite the logistical challenges faced, predominantly relating to the distribution of ground control, survey adaptations detailed in both Chapters 5 and 6 have minimized the impact of systematic errors. Following the suggestions of James and Robson (2014), Javernick et al. (2014) and Dietrich (2016) a convergent camera network has been used to improve the accuracy of initial bundle adjustment to minimize the reliance on GCPs. The process has also been shown to be dependent upon the availability of camera positional data to reduce the computational burden and enable reconstruction as one photogrammetric block to prevent the accumulation of errors when combining multiple blocks. The combination of a convergent camera network and use of a camera GPS has resulted in a reduction in ICP MAE from 0.44 m to 0.23 m and more importantly ICP maximum error from 2. 67 m to 0.63 m between the 2014 and 2015 surveys, respectively.

Objective B) To develop an error model to quantify the quality of SfM data products for modelling fluvial systems.

Chapter 7 presented a conservative approach to quantify localised random errors generated by software variability, specifically a result of the bundle adjustment and dense cloud production. A spatially variable minLOD derived from propagating precision estimates for different surface classes and applying a user defined confidence interval is implemented. Detection of inundated channels and tall vegetated surfaces is shown to be highly unreliable, with the former requiring removal from change detection analysis. In contrast the values obtained for dry-dry non-vegetated surfaces are encouraging and comparable to minLODs used elsewhere in the literature (Brasington et al., 2003; Lane et al., 2003). Future work would benefit from the development of an error model to account for the systematic errors experienced within landscape scale SfM surveys.

Objective C) To assess the evolution of the Dart River from Slipstream to the Wakatipu delta between 2014 and 2015.

The results from Chapters 4 and 6 allude to significant systematic errors between the 2014 and 2015 survey, latterly confirmed through a sensitivity analysis of the reach-wide DoD in Chapter 7. The significant difference in surveys is a result of both the imaging strategies and the distribution of ground control, limiting the change detection analysis to an 11 km downstream reach proximal to the delta. This region has been subject to an assessment of experiments to empirically tie the surveys together through the identification of presumed stable regions. Results provided average positive and negative depth changes of 0.5 m with what was representative of bar top morphologies in the field. Unfortunately, a lack of validation data limits the extent to which inferences can be made, whilst the masking of the channel network from the analysis removes a major driver of morphological change.

9.2. Future Opportunities

The use of SfM for future studies over large fluvial systems is supported providing the following criteria are met: (1) the use of convergent camera networks; (2) the use of camera GPS; (3) consistent GCP distribution between surveys; (4) construction as one photogrammetric block; and (5) the use of original image resolution. The adoption of these strategies will not eliminate systematic errors, however it will mitigate them to enable the use of simple error modelling strategies to undertake change detection analysis. There exists the potential for the accuracy of future surveys to be further refined through, for example, improved sensors that can mitigate the impact of helicopter flight strategies on pixel size and sharpness. Smith et al. (2016) have also highlighted the potential for RTK systems to one day remove the requirement of ground control.

The true potential of SfM and future work lies in taking advantage of the low cost of application and high spatial and temporal frequencies achievable. The potential of SfM to provide repeat analysis to both surveys and interpretation of geomorphological change has been verified across a number of studies (Westoby et al., 2012; Fonstad et al., 2013; Javernick et al., 2014, Smith et al., 2016). Traditional surveying techniques such as LiDAR and TLS remain constrained by cost and practicality issues that limit their frequency of deployment. The subsequent infrequency of surveying can act to restrict our understanding of geomorphological change, as processes are averaged over time (Rosser et al., 2017). In contrast, the SfM pipeline provides the chance to assess landscape scale evolution processes over previously unrealised time periods, giving the opportunity to expand the area of interest further to encompass valley sides and enable the capture of hill slope processes concurrently with channel morphology. Alongside the increased temporal frequency, this could provide a 4D dataset of 'live' hill slope channel coupling, whereby specific hill slope events may be correlated to specific topographic evolution of the channel. This would also provide a magnitude range of sediment input and channel response scenarios to further assess sediment wave propagation theory.

On a smaller scale and perhaps most fittingly, considering the genesis of Bundler as a means for using online sourced imagery (Snavely et al., 2006), there is great potential for crowd sourcing imagery of specific areas of interest. For example, the Slipstream

debris fan may be viewed from across the channel on the Dart-Rees track. Walkers may be encouraged to take photography from a number of specifically established viewing points, chosen to represent a convergent imaging array of the debris fan along the track. These may be uploaded to an online repository and then grouped based on date of imagery to enable the remote reconstruction of the debris fan. Whilst rudimentary in terms of pure accuracy, the increased temporal frequency and spatial distribution offered by this approach at minimal cost is highly attractive.

Whilst there is significant potential for crowd sourcing SfM imagery, there remain the same caveats that have been present throughout this thesis, namely quality control. A number of measures may be taken to maximise consistency, these are:

(1) Maintain a sparse network of ground control that is re-surveyed on a consistent basis. The frequency would be dependent on the feature and active nature of the terrain.
(2) Designate image locations that provide an optimum convergent network and ensure consistency.
(3) Time stamped imagery to ensure that models are constructed only from images within user specified time periods – again depending on the nature of the terrain and weather conditions.
(4) A final degree of QC exercised when selecting imagery to remove that which is blurred, suffers from poor lighting or is particularly low resolution.

Any crowd sourced project would require a trial period to determine the consistency of model reconstruction against a recognised SfM survey protocol. A simple DoD analysis at a number of user defined dates would provide an insight into the level of accuracy achieved. This may also be used to better inform the frequency at which GCP's must be re-surveyed, for example, by tracking model deterioration over time in relation to re-survey dates.

9.3. Summary

This thesis has performed the first SfM reconstruction of a braided river at the landscape-scale, undertaken as one photogrammetric block, consisting of a convergent imaging network. Despite the presence of, at times, unexpected significant random and

systematic errors, this thesis has sought to provide solutions and an approach to mitigate these through methodological adjustments. The quality of the reconstructions in 2015, as measured by ICP measurements, has indicated the potential for this process to provide the means to undertake landscape-scale change detection in the future. Initial attempts to perform this have resulted in a DoD at an unprecedented scale for an 11 km reach of the Dart River, with estimates of geomorphological change broadly consistent with expectations.

10. Bibliography

Abdel-Aziz, Y. and Karara, H. (1971) Direct Linear Transformation from Comparator Coordinates into Object Space Coordinates in Close-Range Photogrammetry. *Papers from the American Society of Photogrammetry, Symposium on Close-Range Photogrammetry* 433: 1–18.

Agisoft; Agisoft PhotoScan User Manual Professional Edition, Version 1.3 (2017) [online] Available at: http://www.agisoft.com/pdf/photoscan-pro_1_3_en.pdf [Accessed 15/09/2017]

Ashmore, P. E. (1991) Channel Morphology and Bed Load Pulses in Braided, Gravel-Bed Streams. *Geografiska Annaler. Series A, Physical Geography.* 73(1): 37–52.

Baily, B., Collier, P., Farres, P., Inkpen, R. and Pearson, A. (2003) Comparative assessment of analytical and digital photogrammetric methods in the construction of DEMs of geomorphological forms. *Earth Surface Processes and Landforms*. 28(3): 307–320.

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: 2813–2837.

Bay, H., Ess, A. and Tuytelaars, T. (2008) Speeded-Up Robust Features (SURF). *Computer Vision and Image Understanding* 110(3): 346–359.

Behan, A. (2000) On the matching accuracy of rasterised scanning laser altimeter data. *International Archives of Photogrammetry and Remote Sensing.* 33(2): 75-80

Besl, P. J and. McKay, N. D. (1992) A Method for Registration of 3-D Shapes. *IEEE Trans. on Pattern Analysis and Machine Intelligence*. Los Alamitos, CAUSA: IEEE Computer Society. 14 (2): 239–256

Bertin, S., Friedrich, H., Delmas P., Chan, E. and Gimel'farb, G. (2015) Digital stereo photogrammetry for grain-scale monitoring of fluvial surfaces: Error evaluation and workflow optimisation. *ISPRS Journal of Photogrammetry and Remote Sensing* 101: 193–208.

Bethmann, F. and Luhmann, T. (2015) Semi-Global Matching in Object Space. *ISPRS*

Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences 40(3): 23–30.

Bowen, L. and Danya, Y. (2014) Calculation of vehicle real-time position overcoming the GPS positioning latency with MEMS INS", Proceedings of 2014 IEEE International Conference on Service Operations and Logistics, and Informatics.

Brasington, J., Rumsby, D. T. and McVey, R. A. (2000) Monitoring and modelling morphological change in a braided gravel-bed river using high resolution GPS-based survey. *Earth Surface Processes and Landforms* 25(9): 973–990.

Brasington, J., Langham, J. and Rumsby, B. (2003) Methodological sensitivity of morphometric estimates of coarse fluvial sediment transport. *Geomorphology* 53(3–4): 299–316.

Brasington, J., Vericat, D. and Rychkov, I. (2012) Modeling river bed morphology, roughness, and surface sedimentology using high resolution terrestrial laser scanning. *Water Resources Research* 48(11)

Bretar, F., Arab-Sedze, M., Champion J., Pierrot-Deseilligny, M., Heggy, E. and Jacquemoud, S. (2013) An advanced photogrammetric method to measure surface roughness: Application to volcanic terrains in the Piton de la Fournaise, Reunion Island. *Remote Sensing of Environment* 135: 1–11.

Brodu, N. and Lague, D. (2012) 3D terrestrial LiDAR data classification of complex natural scenes using a multi-scale dimensionality criterion: Applications in geomorphology. *ISPRS Journal of Photogrammetry and Remote Sensing* 68: 121–134.

Brown, D. C. (1966) Decentring distortion of lenses. *Photogrammetric Engineering* 32: 444–462.

Brunier, G., Fleury, J., Anthony, E, J., Pothin, V. Vella, C., Dussouilez, P., Gardel, A. and Michaud, E. (2016) Structure-from-Motion photogrammetry for high-resolution coastal and fluvial geomorphological surveys. *Géomorphologie: Relief, Processus, Environnement* 22(2): 147–161.

Cao, Z. and Carling, P. A. (2003) On evolution of bed material waves in alluvial rivers.

Earth Surface Processes and Landforms, John Wiley and Sons, Ltd. 28(4): 437–441.

Chai, T. and Draxler, R. R. (2014) Root Mean Square Error (RMSE) or Mean Absolute Error (MAE)?—Arguments against Avoiding RMSE in the Literature. *Geoscientific Model Development* 7: 1247–1250.

Chandler, J. (1999) Effective application of automated digital photogrammetry for geomorphological research. *Earth Surface Processes and Landforms* 24(1): 51–63.

Civera, J., Grasa, O. G., Davison, A. J. and Montiel, J. M. M. (2009) 1-point RANSAC for EKF-based structure from motion. In: *2009 IEEE/RSJ International Conference on Intelligent Robots and Systems, IROS 2009*, pp. 3498–3504.

Clapuyt, F. Vanacker, V. and Van Oost, K. (2016) Reproducibility of UAV-based earth topography reconstructions based on Structure-from-Motion algorithms. *Geomorphology*.

Collier, P. (2002) The impact on topographic mapping of developments in land and air survey: 1900–1939. *Cartography and Geographic Information Science* 29(3): 155–174.

Cox, S. C., Rattenbury, M. S., McSaveney, M. J. and Hamling, I. J. (2014) *Activity of the landslide Te Horo and Te Koroka fan, Dart River, New Zealand during January 2014*. Lower Hutt, N.Z.

Cui, Y. and Parker, G.(2003) Sediment pulses in mountain rivers: 2. Comparison between experiments and numerical predictions. *Water Resources Research* 39(9): 1–11.

Cui, Y. and Parker, G. (2005) Numerical Model of Sediment Pulses and Sediment-Supply Disturbances in Mountain Rivers. *Journal of Hydraulic Engineering* 131(8): 646–656.

Cunliffe, A. M., Anderson, K., DeBell, L. and Duffy, J. P. A. (2017) A UK Civil Aviation Authority (CAA)-approved operations manual for safe deployment of lightweight drones in research. *International Journal of Remote Sensing* 38(8–10): 2737–2744.

D'Oleire-Oltmanns, S., Marzolff, I., Peter, K. D. and Ries, J. B. (2012) Unmanned aerial vehicle (UAV) for monitoring soil erosion in Morocco. *Remote Sensing* 4(11): 3390–3416.

Dall'Asta, E. and Roncella, R. (2014) A comparison of semiglobal and local dense matching algorithms for surface reconstruction. *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences - ISPRS Archives* 40(5): 187–194.

Davies, T. R., Manville, V., Kunz, M. and Donadini, L. (2007) Modeling Landslide Dambreak Flood Magnitudes: Case Study. *Journal of Hydraulic Engineering* 133(7): 713–720.

Dewitt, B. A. and Wolf, R. (2000) Elements of Photogrammetry(with Applications in GIS), McGraw-Hill Higher Education

Dietrich, J. T. (2016) Riverscape mapping with helicopter-based Structure-from-Motion photogrammetry. *Geomorphology* 252: 144–157.

Dietrich, J. T. (2017) Bathymetric Structure-from-Motion: extracting shallow stream bathymetry from multi-view stereo photogrammetry. *Earth Surface Processes and Landforms* 42(2): 355–364.

Earlie, C. S., Masselink, G., Russell, P., Shail, R. and Kingston, K. (2013) Accurate 3D point cloud comparison and volumetric change analysis of Terrestrial Laser Scan data in a hard rock coastal cliff environment. In: *American Geophysical Union, Fall Meeting 2013*.

Eltner, A., Kaiser, A., Abellan, A., and Schindewolf, M. (2017) Time lapse structure from-motion photogrammetry for continuous geomorphological monitoring. *Earth Surf. Process. Landforms.* doi: 10.1002/esp.4178.

Eltner, A. and Schneider, D. (2015) Analysis of different methods for 3-D reconstruction of natural surfaces from parallel-axes UAV images, *Photogramm. Rec*. 30(151):279-299

Filin S. (2003) Recovery of systematic biases ilaser altimetry data using natural surfaces. *ISPRS Journal of Photogrammetric Engineering and Remote Sensing*. 69: 1235-1242.

Fisher, A. T. (1998) Permeability within basaltic oceanic crust. *Reviews of Geophysics* 36(2): 143–182.

Fonstad, M. A., Dietrich, J. T., Courville. B. C., Jensen, J. L. and Carbonneau, P. E. (2013) Topographic structure from motion: A new development in photogrammetric measurement. *Earth Surface Processes and Landforms*. 38: 421-430

Fryer, J. G. and Mitchell, H. L. (1987) Radial distortion and close-range stereophotogrammetry. *Aust. J. Geod. Photogram. Surv.* (March): 123–138.

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

Furukawa, Y., Curless, B., Szeliski, S, M. and Seitz, R. (2010) Towards Internet-scale multi-view stereo. In: *IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, San Francisco, CA, pp. 1434–1441.

Galliani, S. and Schindler, K. (2015) Massively Parallel Multiview Stereopsis by Surface Normal Diffusion. *International Conference on Computer Vision (ICCV)*: 873–881.

Gilbert, G. K. (1917) *Hydraulic mining debris in the Sierra Nevada*. *U.S. Geological Survey Professional Paper 105*.

Giménez, R., Marzolff, I. and Campo, M, A., Seeger, M., Ries, J. B., Casali, J. and Alvarez-Mozos, J. (2009) Accuracy of high-resolution photogrammetric measurements of gullies with contrasting morphology. *Earth Surface Processes and Landforms* 34(14): 1915–1926.

Gneeniss, A. S., Mills, J. P. and Miller, P. E. (2015) In-flight photogrammetric camera calibration and validation via complementary LiDAR. *ISPRS Journal of Photogrammetry and Remote Sensing* 100: 3–13.

Gómez-Gutiérrez, Á., de Sanjosé-Blasco, J. J. ,de Matías-Bejarano, J. and Berenguer-Sempere, F. (2014a) Comparing Two Photo-Reconstruction Methods to Produce High Density Point Clouds and DEMs in the Corral del Veleta Rock Glacier (Sierra Nevada, Spain). *Remote Sensing* 6(6): 5407.

Gómez-Gutiérrez, Á., Schnabel, S., Berenguer-Sempere, F., Lavado-Contador, F. and Rubio-Delgado, J. (2014b) Using 3D photo-reconstruction methods to estimate gully headcut erosion. *CATENA* 120: 91–101.

Govender, N. (2009) Evaluation of Feature Detection Algorithms for Structure from Motion. *Csir*.

Griffiths, G. A. (1993) Sediment Translation Waves in Braided Gravel-Bed Rivers. *Journal of Hydraulic Engineering*. 119: 924-937

Gruen, A., and Akca, D. (2005) Least squares 3D surface and curve matching. *ISPR Journal of Photogrammetry and Remote Sensing.* 59: 151-174.

Haala, N. and Rothermel, M. (2012) Dense multi-stereo matching for high quality digital elevation models. *Photogrammetrie-Fernerkundung-Geoinformation* 4: 331–343.

Harris, C. and Stephens, M. (1988) A combined corner and edge detector. In: *Fourth Alvey Vision Conference*, pp. 10–5244.

Hartley, R. and Zisserman, A. (2004) *Multiple view geometry in computer vision*. 2nd ed. Cambridge: Cambridge University Press.

Harwin, S. and Lucieer, A. (2012) Assessing the accuracy of georeferenced point clouds produced via multi-view stereopsis from Unmanned Aerial Vehicle (UAV) imagery. *Remote Sensing* 4(6): 1573–1599.

Harwin, S., Lucieer, A. and Osborn, J. (2015) The Impact of the Calibration Method on the Accuracy of Point Clouds Derived Using Unmanned Aerial Vehicle Multi-View Stereopsis. *Remote Sensing*, Multidisciplinary Digital Publishing Institute 7(9): 11933–11953.

Hirschmuller, H. (2005) Accurate and Efficient Stereo Processing by Semi-Global Matching and Mutual Information. In: *Proceedings of the 2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05) - Volume 2 - Volume 02*, CVPR '05, Washington, DC, USA: IEEE Computer Society, pp. 807–814.

Hirschmuller, H. (2008) Stereo Processing by Semiglobal Matching and Mutual Information. *IEEE Transactions on Pattern Analysis and Machine Intelligence*.

Hirschmüller, H., Buder, M. and Ernst, I. M. (2012) Memory efficient semi-global matching. *ISPRS annals of the photogrammetry, remote sensing and spatial*

information sciences 1(3): 371–376.

Hudzietz, B. P. and Saripalli, S. (2011) An experimental evaluation of 3D terrain mapping with an autonomous helicopter. *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences* 38: 137–142.

Hugenholtz, C. H., Whitehead, K., Brown, O. W., Barchyn, T. E., Moorman, B. J., LeClair, A., Riddell, K. and Hamilton, T. (2013) Geomorphological mapping with a small unmanned aircraft system (sUAS): Feature detection and accuracy assessment of a photogrammetrically-derived digital terrain model. *Geomorphology* 194: 16–24.

Huggel, C., Clague, J. J. and Korup, O. (2012) Is climate change responsible for changing landslide activity in high mountains? *Earth Surface Processes and Landforms*, 37(1): 77–91.

Immerzeel, W. W., Kraaijenbrink, P. D. A., Shea, J. M., Shrestha, A. B., Pellicciotti, F., Bierkins, M. F. P. and De Jong, S. M. (2014) High-resolution monitoring of Himalayan glacier dynamics using unmanned aerial vehicles. *Remote Sensing of Environment,* 150: 93–103.

James, L. A. (2010) Secular sediment waves, channel bed waves, and legacy sediment. *Geography Compass*. 4: 576-598.

James, L. A., Hodgson, M. E., Ghoshal, S. and Latiolais, M. M. (2012) Geomorphological change detection using historic maps and DEM differencing: The temporal dimension of geospatial analysis. *Geomorphology* 137(1): 181–198.

James, M. R. and Robson, S. (2012) Straightforward reconstruction of 3D surfaces and topography with a camera: Accuracy and geoscience application. *Journal of Geophysical Research* 117(F3): F03017.

James, M. R. 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, M. R., Ilic, S. and Ruzic, I. (2013) Measuring 3D coastal change with a digital camera. In: *Coastal Dynamics: Bordeaux*, pp. 893–904.

James, M. R., Robson, S. and Smith, M. W. (2017) 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: 1769-1788.

Jaud, M., Passot, S., Le Bivic, R., Delacourt, C., Grandjean, P. and Le Dantec, N. (2016) Assessing the accuracy of high resolution digital surface models computed by PhotoScan and MicMac in sub-optimal survey conditions. *Remote Sensing* 8(6): 465.

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

Javernick, L., Hicks, D. M., Measures, R., Caruso, B. and Brasington, J. (2016) Numerical Modelling of Braided Rivers with Structure-from-Motion-Derived Terrain Models. *River Research and Applications* 32(5): 1071–1081.

Johnson, K., Nissen, E. and Saripalli, S. (2014) Rapid mapping of ultrafine fault zone topography with structure from motion. *Geosphere* 10(5): 969–986.

Kääb, A., Girod, L. and Berthling, I. (2014) Surface kinematics of periglacial sorted circles using structure-from-motion technology. *The Cryosphere*, Copernicus GmbH 8(3): 1041–1056.

Kaiser, A., Neugirg, F., Rock, G., Muller, C., Haas, F., Ries, J. and Schidt, J. (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.

Kang, S.B. and Weiss, R. (1997) Characterization of errors in compositing panoramic images. In IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'97), San Juan, Puerto Rico, pp. 103–109

Korup, O. (2005) Large landslides and their effect on sediment flux in South Westland, New Zealand. *Earth Surface Processes and Landforms*, 30(3): 305–323.

Kumar, R and Hanson, A. R. (1994) Robust Methods for Estimating Pose and a Sensitivity Analysis, In CVGIP: Image Understanding, 60(3): 313-342

Kyto, M., Nuutinen, M. and Oittinen, P. (2011) Method for measuring stereo camera depth accuracy based on stereoscopic vision. *SPIE Proceedings: Three-Dimensional Imaging, Interaction, and Measurement.* 7864

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

Lane, S. N. (2000) The Measurement of River Channel Morphology Using Digital Photogrammetry. *The Photogrammetric Record* 16(96): 937–961.

Lane, S. N., Richards, K. S. and Chandler, J. H. (1994) Developments in monitoring and modelling small-scale river bed topography. *Earth Surface Processes and Landforms*, 19(4): 349–368.

Lane, S. N., Westaway, R. M. and Hicks, M. D. (2003) Estimation of erosion and deposition volumes in a large, gravel-bed, braided river using synoptic remote sensing. *Earth Surface Processes and Landforms*, 28(3): 249–271.

Leary, R. J., Hensleigh, J. W., Wheaton, J. M. and DeMeurichy, K. D. (2012) *Recommended Geomorphological Change Detection Procedures for Repeat TLS Surveys from Hells Canyon, Idaho*. Logan, Utah.

Li, F., Li, E., Shafiee, M. J., Wong, A. and Zelek, J. (2015) Dense Depth Map Reconstruction from Sparse Measurements Using a Multilayer Conditional Random Field Model.

Lisle, T. E., Cui, Y., Parker, G., Pizzuto, J. E. and Dodd, A. M. (2001) The dominance of dispersion in the evolution of bed material waves in gravel-bed rivers. *Earth Surface Processes and Landforms*, 26(13): 1409–1420.

Lowe, D. G. (1999) Object recognition from local scale-invariant features. In: *Proceedings of the International Conference on Computer Vision*, pp. 1150–1157.

Lowe, D. G. (2004) Distinctive Image Features from Scale-Invariant Keypoints. *International Journal of Computer Vision* 60(2): 91–110.

Lucieer, A., de Jong, S. M. and Turner, D. (2014a) Mapping landslide displacements using Structure from Motion (SfM) and image correlation of multi-temporal UAV photography. *Progress in Physical Geography* 38(1): 97–116.

Lucieer, A., Turner, D., King, D. H. and Robinson, S. A. (2014b) Using an Unmanned Aerial Vehicle (UAV) to capture micro-topography of Antarctic moss beds. *International Journal of Applied Earth Observation and Geoinformation* 27: 53–62.

Luhmann, T. (2003) Analytische Auswerteverfahren. Nahbereichsphotogrammetrie – Grundlagen, Methoden und Anwendungen (2. überarbeitete Auflage). Wichmann, Heidelberg, 586 pages: 233–358. MAAS, H.–G., 1

Luhmann, T., Robson, S., Kyle, S. and Boehm, S. (2013) *Close-Range Photogrammetry and 3D Imaging*. Berlin, Boston: De Gruyter.

Madej, M. A. and Ozaki, V. (1996) Channel response to sediment wave propagation and movement, Redwood Creek, California, USA. *Earth Surface Processes and Landforms* 21: 911–927.

Mancini, F., Dubbini, M., Gattelli, M., Stecchi, F., Fabbri, S. and Gabbianeli, S. (2016) Using unmanned aerial vehicles (UAV) for high-resolution reconstruction of topography: The structure from motion approach on coastal environments. *Remote Sensing* 5: 6880–6898.

Marcis, M. (2013) Quality of 3D models generated by SfM technology. *Slovak Journal of Civil Engineering* XXI: 13–24.

Marr, D. and Poggio, T. (1979) A Computational Theory of Human Stereo Vision. *Proceedings of the Royal Society of London B: Biological Sciences*, 204(1156): 301–328.

Marteau B, Vericat D, Gibbins C, . Batalla, R. J. and Green, D. R. (2017) Application of Structure-from-Motion photogrammetry to river restoration. *Earth Surface Processes and Landforms* 42(3): 503–515.

Massey, C. I., Townsend, D. B., Rosser, B. J., Villeneuve, M., McColl, S., Davidson, J., Carey, J. M., Lyndsell, B. M., Lukovic, B., Singeisen, C., Dellow, G. D and Cox, S. C.

(2017) *Landslides generated by the Kaikoura 2016 earthquake. Version 1*.

Mathews, A. J. and Jensen, J. L. R. (2013) Visualizing and Quantifying Vineyard Canopy LAI Using an Unmanned Aerial Vehicle (UAV) Collected High Density Structure from Motion Point Cloud. *Remote Sensing* 5: 2164–2183.

McColl, S. T. and Davies, T. R. (2011) Evidence for a rock-avalanche origin for 'The Hillocks' 'moraine', Otago, New Zealand. *Geomorphology* 127(3): 216–224.

McMillan, H. K. and Brasington, J. (2007) Reduced complexity strategies for modelling urban floodplain inundation. *Geomorphology* 90(3–4): 226–243.

Mellor, J. P., Teller, S. and Lozano-Pérez, T. (1997) Dense Depth Maps from Epipolar Images. *Proceedings of Image Understanding Workshop, New Orleans, Louisiana, USA* (August): 893–900.

Mémoli, F. and Sapiro, G. (2005) Distance Functions and Geodesics on Submanifolds of \R^d and Point Clouds. *SIAM Journal on Applied Mathematics* 65(4): 1227–1260.

Micheletti, N., Chandler, J. H. and Lane, S. N. (2015) Structure from Motion (SfM) Photogrammetry. *Geomorphological Techniques* 2: 1–12.

Milan, D. J., Heritage, G. L. and Hetherington, D. (2007) Application of a 3D laser scanner in the assessment of erosion and deposition volumes and channel change in a proglacial river. *Earth Surface Processes and Landforms*, 32(11): 1657–1674.

Miller, H. J. and Goodchild, M. F. (2015) Data-driven geography. *GeoJournal* 80: 449–461.

Ni, K. and Dellaert, F. (2012) HyperSfM. In: *In Proc. of the IEEE Int. Conf. on 3D Imaging, Modeling, Processing, Visualization and Transmission*.

Niethammer, U., James, M. R., Rothmund, S., Travelletti, J. and Joswig, W. (2012) UAV-based remote of the Super-Sauze landslide: Evaluation and results. *Engineering Geology*, 128: 2–11.

Notebaert, B., Verstraeten, G., Govers, G. and Poesen, J. (2009) Qualitative and quantitative applications of LiDAR imagery in fluvial geomorphology. *Earth Surface*

Processes and Landforms 34(2): 217–231.

NZAM (2011) *QLDC LiDAR Acquisition 11.044. Metadata report,*.

ORC. (2008) Channel morphology and sedimentation in the Rees River. [online] Available at: http://www.orc.govt.nz/Documents/Publications/Research%20And%20Technical/surface-water-quality/Channel%20Morphology%20and%20Sedimentation%20in%20the%20Rees%20River%20October%202008.pdf [Accessed: 12/09/2017]

Pierson, T. C., Pringle, P. T. and Cameron, K. A. (2011) Magnitude and timing of downstream channel aggradation and degradation in response to a dome-building eruption at Mount Hood, Oregon. *Geological Society of America Bulletin* 123(1–2): 3–20.

Rastgar, H. (2013) Robust Self-calibration and Fundamental Matrix Estimation in 3D Computer Vision, Doctoral Dissertation, Ottawa: Univeristy of Ottawa.

Redolfi, M., Tubino, M., Bertoldi, W., Brasington, J. (2016) Analysis of reach-scale elevation distribution in braided rivers: Definition of a new morphologic indicator and estimation of mean quantities. *Water Resources Research* 52(8): 5951–5970.

Robertson, D. P. and Cipolla, R. (2009) Structure from Motion. In: *In: Varga, M., (Ed.), Practical Image Processing and Computer Vision*, New Yor: John Wiley and Sons, Ltd.

Rosnell, T. and Honkavaara, E. (2012) Point cloud generation from aerial image data acquired by a quadrocopter type micro unmanned aerial vehicle and a digital still camera. *Sensors* 12(1): 453–480.

Rosser, N., Williams, J., Hardy, R. and Brain, M. (2017) 4D monitoring of actively failing rockslopes. In 19th EGU General Assembly, EGU2017, Vienna, Austria, p.4789

Rumsby, B. T., Brasington, J., Langham, J. A., McLelland, S. J., Middleton, R. and Rollinson, G. (2008) Monitoring and modelling particle and reach-scale

morphological change in gravel-bed rivers: Applications and challenges. *Geomorphology*. 93(1-2): 40-54.

Rupnik, E., Nex, F., Toschi, I. and Remondino, F. (2015) Aerial multi-camera systems: Accuracy and block triangulation issues. *ISPRS Journal of Photogrammetry and Remote Sensing* 101: 233–246.

Ružić, I., Marović, I., Benac, Č. and Ilic, S. (2014) Coastal cliff geometry derived from structure-from-motion photogrammetry at Stara Ba{š}ka, Krk Island, Croatia. *Geo-Marine Letters* 34(6): 555–565.

Sangireddy, H., Stark, C. P. and Passalacqua, P. (2017) Multiresolution analysis of characteristic length scales with high-resolution topographic data. *Journal of Geophysical Research: Earth Surface* 122(7): 1296–1324.

Scharstein, D. and Szeliski, R. (2002) A Taxonomy and Evaluation of Dense Two-Frame Stereo Correspondence Algorithms. *International Journal of Computer Vision* 47(1): 7–42.

Shan, J. and Toth, C. K. (2009) *Topographic laser ranging and scanning*. Boca Raton, USA: CRC Press.

Sirguey, P. (2015) *RPAS Operation - Hillocks, Dart River. Unpublished report 15-004.001.R1.0*.

Smith, M. W., Carrivick, D. J. and Quincey, D. J. (2016) Structure from motion photogrammetry in physical geography. *Progress in Physical Geography* 40(2): 247–275.

Smith, M. W. (2014) Roughness in the Earth Sciences. *Earth-Science Reviews* 136: 202–225.

Smith, M. W. 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.

Smith, M. W., Carrivick, J. L., Hooke, J. and Kirby, M. J. (2014) Reconstructing flash flood

magnitudes using 'Structure-from-Motion': A rapid assessment tool. *Journal of Hydrology* 519: 1914–1927.

Snapir, B., Hobbs, S. and Waine, T. W. (2014) Roughness measurements over an agricultural soil surface with Structure from Motion. *ISPRS Journal of Photogrammetry and Remote Sensing*, International Society for Photogrammetry and Remote Sensing, Inc. (ISPRS) 96: 210–223.

Snavely, N., Seitz, S. M. and Szeliski, R. (2006) Photo tourism: exploring photo collections in 3D. *ACM Transactions on Graphics (TOG)* 25(3): 835–846.

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

Solbø, S. and Storvold, R. (2013) Mapping Svalbard glaciers with the cryowing UAS. In: *ISPRS International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, XL-1/W2*, pp. 373–377.

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

Sutherland, D. G., Ball, M. H., Hilton, S. J. and Lisle, T. E. (2002) Evolution of a landslide-induced sediment wave in the Navarro River, California. *Bulletin of the Geological Society of America* 114(8): 1036–1048.

Sutherland, R., Eberhart-Phillips, D., Harris, R. A., Stern, T., Beavan, J., Eliis, S., Henry, S., Cox, S. J., Norris, R. J., Berryman, K. R., Townsend, J., Bannister, S., Pettinga, J., Leitner, B., Wallace, L., Little, T. A., Cooper, A. F., Yetton, M. and Stirling, M. (2006) Do great earthquakes occur on the Alpine fault in central South Island, New Zealand? *Geophysical Monograph, SIGHT*.

Tamminga, A. D., Eaton, B. C. and Hugenholtz, C. H. (2015) UAS-based remote sensing of fluvial change following an extreme flood event. *Earth Surface Processes and Landforms* 40(11): 1464–1476.

Tarolli, P. (2014a) High-resolution topography for understanding Earth surface

processes: Opportunities and challenges. *Geomorphology*, 216: 295–312.

Taylor, R. (1982) *An Introduction to Error Analysis: The Study of Uncertainties in Physical Measurements*. University Science Books.

Tonkin, T. N., Midgley, N. G., Graham, D. J. and Labadz, J. C. (2014) The potential of small unmanned aircraft systems and structure-from-motion for topographic surveys: A test of emerging integrated approaches at Cwm Idwal, North Wales. *Geomorphology* 226: 35–43.

Triggs, B., McLauchlan, P. F., Hartley, R. I. and Fitzgibbo, F. W. (2000) Bundle Adjustment --- A Modern Synthesis. In: Triggs B, Zisserman A, and Szeliski R (eds), *Vision Algorithms: Theory and Practice: International Workshop on Vision Algorithms Corfu, Greece, September 21--22, 1999 Proceedings*, Berlin, Heidelberg: Springer Berlin Heidelberg, pp. 298–372.

Turnbull, I. M. (2000) Geology of the Wakatipu area. Institute for Geological and Nuclear Sciences 1:250 000 geological map 18. 1 sheet _ 72 p. Institute of Geological and Nuclear Sciences Limited, Lower Hutt, New Zealand.

Turner, D., Lucieer, A. and Watson, C. (2012) An automated technique for generating georectified mosaics from ultra-high resolution Unmanned Aerial Vehicle (UAV) imagery, based on Structure from Motion (SFM) point clouds. *Remote Sensing* 4(5): 1392–1410.

Vaaja, M., Kukko, A., Kaartinen, H., Kurkela, M., Kasvi, E., Flener, C., Hyyppa, H., Hyyppa, J., Jarvela, J. and Alho. P. (2013) Data processing and quality evaluation of a boat-based mobile laser scanning system. *Sensors* 13(9): 12497–12515.

van Lekkerkerk, H. (2017) *GNSS Survey and Engineering*. Geomares Publishing.

Verhoeven, G. (2011) Taking computer vision aloft - archaeological three-dimensional reconstructions from aerial photographs with photoscan. *Archaeological Prospection*, 18(1): 67–73.

Verhoeven, G., Doneus, M., Briese, C. and Vermeulen, F. (2012) Mapping by matching: a computer vision-based approach to fast and accurate georeferencing of

archaeological aerial photographs. *Journal of Archaeological Science* 39(7): 2060–2070.

Verhoeven, G. J. J. (2009) Providing an archaeological bird's-eye view â€ ' an overall picture of ground-based means to execute low-altitude aerial photography (LAAP) in Archaeology. *Archaeological Prospection*, John Wiley and Sons, Ltd. 16(4): 233–249.

Vericat, D., Smith, M. W. and Brasington, J. (2014) Patterns of topographic change in sub-humid badlands determined by high resolution multi-temporal topographic surveys. *CATENA* 120: 164–176.

Vericat, D., Wheaton, J. M. and Brasington, J. (2017) Revisiting the Morphological Approach: Opportunities and Challenges with Repeat High-Resolution Topography. In: *Gravel-Bed Rivers*, John Wiley and Sons, Ltd, pp. 121–158.

Wackrow, R. and Chandler, J. H. (2008) A convergent image configuration for DEM extraction that minimizes the systematic effects caused by an inaccurate lens model. *Photogrammetric Record*, 23: 6-18.

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

Wackrow, R., Chandler, J. H. and Bryan, P. (2007) Geometric consistency and stability of consumer-grade digital cameras for accurate spatial measurement. *Photogrammetric Record*, 22: 121-134.

Westoby, M. J., Brasington, J., Glasser, N. F., Hambrey, J. and Reynolds, J. M. (2012) 'Structure-from-Motion' photogrammetry: A low-cost, effective tool for geoscience applications. *Geomorphology*, 3: 171-199.

Westway, R. M., Lane, S. N. and Hicks, D. M. (2003) Remote survey of large-scale braided, gravel-bed rivers using digital photogrammetry and image analysis, *International Journal of Remote Sensing*. 24(4): 795-815.

Wheaton, J. M., Brasington, J., Darby, S. E, and Sear, D. A. (2010) Accounting for

uncertainty in DEMs from repeat topographic surveys: improved sediment budgets. *Earth Surface Processes and Landforms,* 35: 136-156.

Wheaton, J. M., Brasington, J., Darby, S. E., Kasprak, A., Sear, D. A. and Vericat, D. (2013) Morphodynamic signatures of braiding mechanisms as expressed through change in sediment storage in a gravel-bed river. *Journal of Geophysical Research: Earth Surface* 118(2): 759–779.

Wild, M. A. (2012) Growth dynamics of braided gravel-bed river deltas in New Zealand. (Unpublished doctoral dissertation). University of Canterbury, New Zealand.

Williams, R. (2012) DEMs of difference. Online Edi. In: Cook SJ, Clarke LE, and Nield JM (eds), *Geomorphological Techniques*, London, UK: British Society for Geomorphology.

Williams, R., Brasington, J., Vericat, D., Hicks, M., Labrosse, F., Neal, M. (2011) Chapter Twenty – Monitoring Braided River Change Using Terrestrial Laser Scanning and Optical Bathymetric Mapping. In: *Developments in Earth Surface Processes*, pp. 507–532.

Williams, R., Measures, R., Hicks, D. and Brasington, J. (2016a) Assessment of a numerical model to reproduce event-scale erosion and deposition distributions in a braided river. *Water Resources Research* 52(8): 6621–6642.

Williams, R. D., Brasington, J., Hicks, M., Measures, R., Rennie, C. D., Vericat, D. (2013) Hydraulic validation of two-dimensional simulations of braided river flow with spatially continuous aDcp data. *Water Resources Research* 49(9): 5183–5205.

Williams, R. D., Brasington, J. and Hicks, D. M. (2016b) Numerical modelling of braided river morphodynamics: Review and future challenges. *Geography Compass* 10(3): 102–127.

Williams, R. D., Tooth, S. and Gibson, M. (2017) The sky is the limit: reconstructing physical geography from an aerial perspective. *Journal of Geography in Higher Education* 41(1): 134–146.

Wolf, P. R and Dewitt, B. A. (1996) Elements of photogrammetry with applications in GIS

(3rd ed.) San Francisco, United States: McGraw-Hill Education.

Woodget, A. S., Carbonneau, P. E., Visser, F., Maddock, I. P. (2015) Quantifying submerged fluvial topography using hyperspatial resolution UAS imagery and structure from motion photogrammetry. *Earth Surface Processes and Landforms*, 40: 47-64.

Xiang, T. and Cheong, L. F. (2003) Understanding the Behavior of SFM Algorithms: A Geometric Approach. *International Journal of Computer Vision* 51(2): 111–137.

Zhen, Y., Liu, X. and Wang, M. (2013) Precise Fundamental Matrix Estimation Based on Inlier Distribution Constraint. In: Lu W, Cai G, Liu W, et al. (eds), *Proceedings of the 2012 International Conference on Information Technology and Software Engineering: Information Technology {and} Computing Intelligence*, Berlin, Heidelberg: Springer Berlin Heidelberg, pp. 237–244.