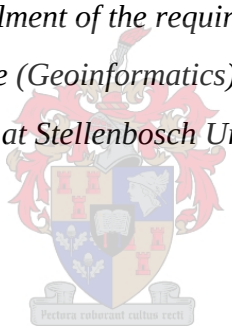
EVALUATION OF SFM AGAINST TRADITIONAL STEREOPHOTOGRAMMETRY AND LIDAR TECHNIQUES FOR DSM CREATION IN VARIOUS LAND COVER AREAS

By

SEBASTIAN MICHAEL BENEDICT WENGER

Thesis submitted in fulfilment of the requirements for the degree of Master of Science (Geoinformatics) in the Faculty of Science at Stellenbosch University



Supervisor: Dr Jaco Kemp

December 2016

DECLARATION

By submitting this report electronically, I declare that the entirety of the work contained therein is my own, original work, that I am the sole author thereof (save to the extent explicitly otherwise stated), that reproduction and publication thereof by Stellenbosch University will not infringe any third party rights and that I have not previously in its entirety or in part submitted it for obtaining any qualification.

December 2016

SUMMARY

Structure from Motion (SfM) is potentially an effective tool for the creation of digital surface models (DSMs) and ortho-mosaiced imagery. The increase in camera resolution and the emergence of affordable and commercial unmanned aerial vehicle (UAV) technology has provided a basis from which SfM can be deployed effectively for DSM creation. Recent studies with SfM have predicted a vertical accuracy of less than a meter in specific study environments.

A need exists for surface models for topographical and engineering surveying but there is limited access to high resolution digital elevation models (DEMs) for parts of South Africa only and low resolution DEMs for the remainder. Potentially SfM could help to bridge this gap, if consistent accuracy can be proven.

To better understand the effectiveness of SfM for Geographic Information Systems (GIS) application, this study compared spatial accuracies of SfM to well understood commercial techniques (traditional stereophotogrammetry and Light Detection and Ranging or LiDAR) in DSM creation in varying land use and land cover environments. This assessment was made by comparing each technique to an independent Global Positioning System (GPS) reference as well as relative to each other.

It was found that LiDAR was the best method for surface reconstruction picking up the highest level of detail and overall the lowest margin of error over different land uses (Root Mean Square Error, RMSE, 0.57m). SfM in the same assessment found an overall RMSE of 0.67m with stereophotogrammetry performing least well with and RMSE of 1.22m. SfM showed limited reconstruction ability in vegetated environments with only 58.2% included within 1m from the LiDAR surface. SfM did, however, pick up formal crop rows and demonstrated a low mean error from a LiDAR surface reference in the vineyard land cover class. Stereophotogrammetry failed to pick up vine rows and displayed higher levels of overgeneralization in vegetated areas with a lower level of inclusion. SfM performed better in built environments with 73.4-81.4% inclusion within 1m from the LiDAR surface and a RMSE range of 028-0.62m from a GPS reference. The inter-DSM agreement of SFM in built environments suggested better propensity for data fusion than stereophotogrammetry. The ability to reconstruct built features was adequate but limited by highly clustered buildings resulting in overgeneralized surfaces.

LiDAR is still the best technique for DSM construction but experiences limitations related to availability, cost, specialist equipment and knowledge. For image-based techniques, SfM out performed stereophotogrammetry but this does not necessarily imply stereophotogrammetry is

redundant as a technique. Instead SfM should be considered the natural extension of photogrammetry for use in smaller, highly detailed and potentially highly temporal environments.

KEY WORDS

SfM, LiDAR, Traditional Photogrammetry, Accuracy Assessment, Digital Surface Model

OPSOMMING

Struktuur-deur-Beweging (SdB) is moontlik ’n effektiewe hulpmiddel om digitale oppervlak modelle (DOMs) en orto-mosaïek beelde te skep. Die toename in kamera-resolusie asook die ontwikkeling van bekostigbare en kommersiële onbemande vliegtuie (OVs) tegnologie het 'n basis geskep vir die effektiewe gebruik van SdB om ’n DOM te skep. Onlangse studies wat verband hou met SdB het ’n vertikale akkuraatheid van minder as ’n meter in spesifieke studieomgewings voorspel. Daar is ’n behoefte aan oppervlak modelle wat vir topologiese- en ingenieursopnames gebruik kan word, maar toegang tot hoë resolusie digitale oppervlak modelle (DOMs) is beperk vir gebiede van Suid-Afrika en slegs lae resolusie DOMs is beskikbaar vir die res van die land. SdB kan potensieel die gaping oorbrug indien konsekwente akkuraathede verkry kan word. Om die effektiwiteit van SdB vir Geografiese Inligting Stelsels (GIS) beter te verstaan het hierdie studie die ruimtelike akkuraathede van SdB met gevestigde, tegnieke (tradisionele stereo-fotogrammetrie en Ligopsporing en afstandskatting of LiDAR) vir DOM-skepping in wisselende grondgebruik- en grondbedekkingsomgewings. Hierdie assessering is aangepak deur elke tegniek met onafhanklike globale posisioneringstelsel (GPS) verwysingspunte, asook met mekaar te vergelyk.

Daar is bevind dal LiDAR die beste metode vir oppervlak herkonstruksie is en die meeste detail vasgevang het asook algeheel die laagste relatiewe fouttelling oor verskillende landgebruike gehad (wortel gemiddelde kwadratiese fout, WGKF, 0.57m). SdB het ’n algehele WGKF van 0.67m behaal, terwyl stereo-fotogrammetrie die swakste WGKF van 1.22m gelewer het. SdB het beperkte vermoë getoon om in begroeide omgewings oppervlaktes te herkonstruktureer en slegs 58.2% is binne 1m van die LiDAR oppervlak. SdB het egter rye in wingerde geïdentifiseer en ’n lae gemiddelde fout getoon toe dit met ‘n LiDAR verwysingsoppervlak vergelyk is. Stereo-fotogrammetrie nie die wingerdrye geïdentifiseer nie, het begroeide gebiede te veel veralgemeen en 'n laer vlak van insluiting getoon. SdB het beter resultate gelewer in beboude gebiede met 73.4-81.4% wat binne 1 m van die LiDAR oppervlak val en het ’n WGKF van 028-0.62m van GPS verwysingspunte af gekry. Die inter-DOM ooreenkoms van SdB in beboude omgewings wys daarop dat SdB moontlik beter kan wees vir data smelting as stereo-fotogrammetrie. Die herkonstruksie van beboude gebiede was voldoende, maar is deur baie dig beboude gebiede wat verdere ongewenste veralgemening van oppervlakmodelle veroorsaak.

LiDAR is steeds die beste tegniek vir DOM-skepping, maar beperkings ontstaan as gevolg van die beskikbaarheid, koste, gespesialiseerde toerusting en kundigheid wat benodig word vir die implementering van hierdie tegniek. Vir beeld-gebaseerde tegnieke lewer SDB beter resultate as stereo-fototogrammetie maar dit beteken nie noodwendig dat stereo-fotogrammetie ’n oorbodige

of verouderde tegniek is nie. SdB moet eerder beskou word as die natuurlike vertakking van fotogrammetrie wat in kleiner, hoogs gedetailleerde omgewings wat moontlik hoë temporale veranderlikheid toon gebruik kan word.

TREFWOORDE

Struktuur-deur-beweging, LiDAR, tradisionele fotogrammetrie, akuuraatheidsassessering, Digitale Oppervlak Model

ACKNOWLEDGEMENTS

I sincerely thank:

- The National Research Foundation (NRF)
- Henie Kiesser (Drones Africa)
- Michal Jancosek (Designing Reality)
- Councillor Lubabalo Makeleni and Thomas (Dunoon)
- Jaco Cockart, Edmond Thomas & Denzel (Afrisam)
- Brendon Masters (Much Asphalt)
- Alan Ross (APM Terminals)
- Andrew Melish (Welbeloond Farms)
- Roland Wingeier, Lionel and Yvette (Propet Fibre)
- Anthony Isaacs (CoCT Waste Management)
- Adriaan Kohn (Green Platoon)
- Jonathan Basckin (Pilot and Artist)
- Geoff Dekker & Niyaaz Ramjam (City Maps)
- All members of the CGA
- Khaleed Ballim (Research Assistant)
- Christine Kelbe for all your help with the language editing of this thesis

To all those that helped with this research and my family for their love and support

CONTENTS

TABLES

FIGURES

ACRONYMS AND ABBREVIATIONS

ALS	Aerial laser scanner
CMPMVS	Centre for Machine Perception multi-view stereo reconstruction software
CoCT	City of Cape Town
CMVS	Clustering view for multi-view stereo
CUDA	Compute unified device architecture
DEM	Digital elevation model
DGPS	Differential global positioning systems
DSM	Digital surface model
DTM	Digital terrain model
EXIF	Exchangeable image file format
GCP	Ground control point
GIS	Geographic information systems
GPS	Global positioning systems
LiDAR	Light Detection and Ranging
MSL	Mean sea level
MVS	Multi-view stereo
PMVS	Patch-based multi-view stereo
RMSE	Root mean square error
RPAS	Remotely piloted aircraft system
RPC	Rational polynomial coefficient
SfM	Structure from motion
SIFT	Scale-invariant feature transform
TLS	Terrestrial laser scanner
TPT	Traditional photogrammetric technique(s)
UAV	Unmanned aerial vehicle

CHAPTER 1: INTRODUCTION

This chapter presents the background of the study and the derivation of the research problem, aim and objectives. A concise discussion of the research design and methodology is presented outlining the structure of the thesis.

1.1 BACKGROUND TO THIS STUDY

The advent of commercially available unmanned aerial vehicle (UAV), greater processing power and improvements in camera technology has seen photogrammetry re-emerge with an accurate and cost-effective competitive technology in the form of Structure from Motion (SfM) for digital surface modelling. From its genesis in the field of computer vision (Ullman 1979), SfM has grown into the current algorithms available in commercial and non-commercial software (Snavely et al. 2006; Snavely 2008; Agarwal et al. 2010; Furukawa & Ponce 2010; Kung et al. 2011). SfM has been used for a multitude of structure-based applications to provide high resolution detailed surface models.

UAV platforms have the potential, along with ever-increasing camera resolution, to produce higher quality digital surface models (DSMs) using SfM rather than its stereophotogrammetric counterpart due to the intrinsic spatial resolution from lower flying heights and increased image overlap requirements (James & Robertson 2012). There is a growing need for DSM creation particularly in the commercial sector. Current methods provide geometric representations of an observed surface commonly derived from aerial imagery or Airborne Laser Scanner (ALS) (Brunn & Weidner 1997).

Stereo-imagery and Light Detection and Ranging (LiDAR) point clouds are common datasets, as indicated by the data availability by geospatial vendors and national government organisations (NGOs). The National Geo-spatial Information (NGI), a division of the Department of Rural Development and Land Reform, has used stereo-imagery to create a national digital elevation model (DEM) with the highest available spatial resolutions at 25m x 25m (NGI 2014). This national DEM is intended for use in hydrological analysis, view shed analysis and 3D path profiling, to name a few, but is primarily used by NGI for the orthorectification of aerial imagery and hill shading (NGI 2014). The City of Cape Town (CoCT) Geospatial Services captures high resolution imagery and LiDAR for better quality maps for research, and corporate GIS for

topographical and engineering surveying (City of Cape Town SDI and GIS Department 2015). As part of the service, LiDAR and high resolution imagery is captured for the Cape Town metropolitan area but does not extend past that area. The capture of these datasets are performed every 1-3 years which can create issues for time sensitive research and higher resolution DEM datasets may not be available in the area of study (City of Cape Town SDI and GIS Department 2015). SfM software can create these DEMs and produce orthorectified image mosaics in a semi-automated/automated process (Snavely et al. 2008). If used in conjunction with UAV technology, this spatial and spectral data can be collected in research areas at a potentially lower cost because of non-specialist skill, commercially available cameras and cheaper flying platforms (James & Robertson 2012). As a technique, existing aerial platforms can mobilise SfM with increased overlap of image capture to achieve similar surface products using existing infrastructure (Mueller 2014).

Many, particularly developed, countries, have an existing national databank of high quality DSM images which serve as a valuable research resource. The Netherlands, for example, have LiDAR generated DSMs (Actueel Hoogtebestand Nederland 2015) for the whole country which are freely available. In the South African situation, and even more so in other developing countries, the existing datasets are much less complete and consist of models acquired through differing techniques with variable accuracies. (NGI 2014; City of Cape Town SDI and GIS Department 2015). These may be insufficient for planned research or specific land use data requiring new data acquisition. SfM can be implemented with lower cost, less equipment, less on site expertise and can provide the required and proven level of accuracy as supported by this study. SfM can potentially become an increasingly important tool for accurate and high resolution DSM acquisition where currently unavailable, particularly in resource challenged settings in Africa.

SfM, as a modern photogrammetric technique, has been reviewed in several disciplines due to its research potential. The photo-tourism niche was the platform for early SfM testing and has led to breakthroughs in this field with the ‘4D cities’ and ‘Building Rome in a Day’ projects (Cornelis et al. 2008; Agarwal et al. 2011). These projects suggested the capability to process large datasets with little or no prior knowledge about the images themselves (Kung et al. 2011). In the field of engineering, it has been applied to construction by comparing current progress (as mapped using SfM) against Building Information Models (BIM) to assess progress of building erection projects (Golparvar-Fard et al. 2011). SfM has also been used to research and monitor geomorphological changes with the use of images and the subsequent generation of point clouds with a high level of

accuracy (Lejot et al. 2007; Rumsby et al. 2008; Neithammer et al. 2012). In geoscience, SfM outputs have been found to be comparable to Terrestrial Laser Scanners (TLS) and Aerial Laser Scanners (ALS) with vertical accuracies in the decimeter range without the use of expensive equipment (Sood & Kaur 2012; Westoby et al. 2012; Fonstad et al. 2013). It is thus believed that SfM may provide an accurate and cost effective platform for Digital Surface Model (DSM) creation in multiple land cover environments. Studies performed with SfM call for further investigation into this technique to determine reliability in multiple land forms (Westoby et al. 2012; Fonstad et al. 2013; Green et al. 2014).

Despite good spatial accuracies existing in research, SfM has not been reliably tested and this stifles the commercial confidence and employment of this technique. LiDAR and stereophotogrammetry have been rigorously studied and have confident predictions of the derived accuracy ranges that can be expected. SfM has isolated studies in single specific environments and but not as a conglomerate. These studies are also derived with individual imagery datasets collected with a variety of equipment reducing the effective cross compatibility of results (e.g. Westoby et al. 2012; Fonstad et al. 2013; Green et al. 2014; Johnson et al. 2014; Tonkin et al. 2014). SfM also has its limitations with regard to specular reflective objects, low texture features, image overlap, homogenous features and a strong correlation of surface reconstruction with the spatial resolution of imagery (Photosynth 2005; Snavely et al. 2008; Jancosek & Pajdla 2011). SfM has the potential to produce high quality and detailed DSMs with the technology available. It is essential, however, to evaluate SfM over differing surface environments with existing DSM generation techniques to further gauge the appropriate use and accuracy of this technology.

In the interest of comparing SfM to more traditional and well established techniques, such as photogrammetric image matching and laser scanning technology for DSM creation, proprietary and industry standard software was used to process and evaluate DSMs for all surface reconstruction techniques in this study. This provided a fairer comparison as open source software has its own technical limitations with specific regard to SfM. Problems experienced with open source software for practical deployment include the legal use of intellectual property rights for commercial use (licensing), software support, limitations with regard to their ability to handle large imagery and point location transformations (geo-referencing). The imagery limitations of open source stem from resolution (where imagery has to be downscaled) and dataset size (non-linear time incrimination for processing) (Wu 2011; Favalli et al. 2012).

1.2 PROBLEM FORMULATION

SfM and Mutli-View Stereo (MVS) is better known in the field of Computer Vision and is relatively novel in the study of geographic information systems (GIS) with limited literature on the technique in this field. Along with the advent of commercially available UAVs, it makes SfM-MVS a potentially viable and powerful tool for GIS application based on the portability and the ability of non-specialists to operate the software (James & Robertson 2012). SfM still needs to be assessed in terms of reliability, accuracy and precision in a wide range of environments before it can be considered as an accepted practice in GIS (Fonstad et al. 2012). Since this statement was made in 2012, the industry has evolved, but SfM is still used to lesser degree than traditional photogrammetry and LiDAR due to issues of familiarity and possibly due to higher project specifications.

This study seeks to understand the surface reconstruction quality of SfM in practically comparative study areas as deployment considerations in the context of existing techniques. A rigorous accuracy assessment is required over multiple land forms and land types to better elucidate the major sources of error, which is essential for the adoption of this method in practice (Westoby et al. 2012). The commercial sector requires validation for the use of fine-scale reconstruction of land forms and surface structures before accepting this technique fully (Harwin & Lucieer 2012). There must be an understanding of the accuracy of SfM when compared to other techniques before a GIS practitioner can motivate the appropriate method for application with specific regard to SfM. Identifying these considerations is an important step in utilising this technique appropriately for DSM construction. As of yet, there is outstanding research that needs to be done to inform any recommendation of this technology in multiple area types (Westoby et al. 2012; Fonstad et al. 2013). Comparative research into the accuracy and intrinsic nature of DSM techniques must be studied individually and together to decide on the appropriate use of SfM for purpose specific applications in a given land cover while identifying the possible sources of error (Westoby et al. 2012).

1.3 RESEARCH AIM AND OBJECTIVES

The aim of this research is to compare spatial accuracies of SfM to those from well understood, current commercial techniques (traditional stereo photogrammetry and LiDAR) in DSM creation in varying environments in the form of a case study.

The research objectives are to:

1. Assess the vertical accuracy of DSMs developed by LiDAR, SfM and stereophotogrammetry to a common area and reference.
2. Perform relative assessments in surface construction of models from derived data comparisons.
3. Identify limitations and evaluate suitability for each method in DSM extraction.

1.4 RESEARCH DESIGN AND THESIS STRUCTURE

The methodological nature of this study follows an empirical approach with comparative quantitative measurements and qualitative evaluations. The research design is illustrated in Figure 1.1. The design is divided into five steps to achieve the research aim: In the first step, prior to the start of the study, an appropriate area was selected containing heterogeneous land use and land covers to test dissimilar environments. This area will form the case study for research as other areas may yield differing results. Secondly, for this area, existing data was collected from various vendors and any outstanding data required for SfM, georeferencing and test samples were manually captured. The third step used this data to process DSMs over the study site using each respective surface reconstruction technique: SfM, LiDAR and Stereophotogrammetry. Each technique was represented adequately and without bias, employing modern/conventional methods for all processes.

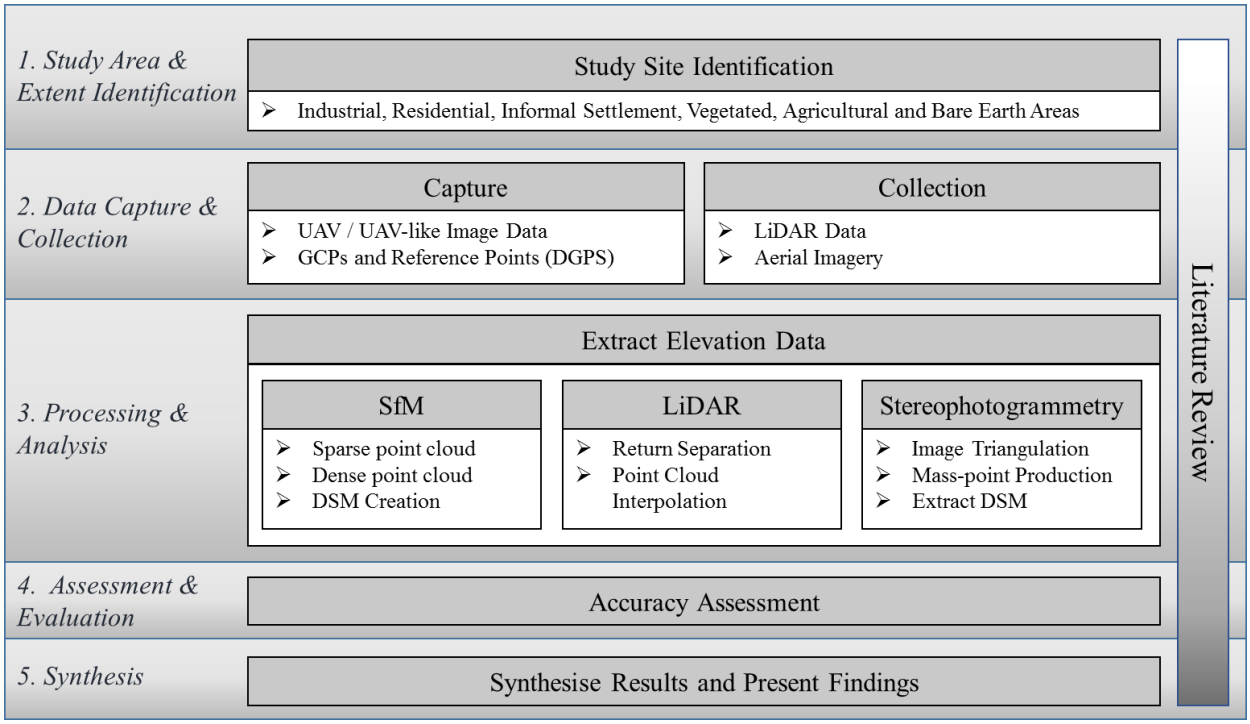


Figure 1.1 The design of research broken up in to five major components for the completion of research objective: 1) Study area and extent identification, 2) Data capture and collection, 3) Processing and analysis, 4) Assessment and evaluation of the data and 5) Synthesis of results.

The fourth step is comprised of multiple experiments including observational descriptions during DSM construction that constitute the accuracy assessment in both an absolute accuracy (quantitative) and relative (qualitative) evaluations. These assessments were broken up into two main types of evaluations: absolute accuracy and relative accuracy. The absolute accuracy was measured vertically by comparing overlapping DSMs of each data capture type (SfM, LiDAR and TPT) to an independent GPS reference. The relative accuracy assessment quantitatively and qualitatively assessed the reconstruction quality of the DSM techniques relative to each other. The fifth and final step, coupled with absolute and relative accuracy results, identified feasibility constraints affecting the creation of these datasets were discussed for practical deployment of SfM to better inform the appropriate technique choice. This workflow was designed to simulate data availability and processing techniques available in the South African context.

The thesis overview is presented below.

- Chapter 1 presents background of SfM and its inclusion into GIS, describing its potential capabilities, and the lack of research governing its use against traditional methods. This leads to the formulation of the aim and objective of the study in relation to this real world issue.
- Chapter 2 is a review of the current literature relating to DSM extraction with LiDAR, SfM and traditional photogrammetric techniques. This chapter includes an overview of case studies that were performed to evaluate the accuracy of SfM.
- Chapter 3 describes the methods used to produce the various DSMs and accuracy assessments to achieve the objectives.
- Chapter 4 presents and interprets the results of this research in terms of both absolute and relative accuracy, with regard to error and variance in DSM construction.
- Chapter 5 summarizes the major findings in relation to the research aim and objectives to form more generalized conclusions. In light of these findings, limitations and recommendations are presented for possible future research.

CHAPTER 2: LITERATURE REVIEW

This chapter reviews the current literature with regard to each tested technique used in this study in terms of its functionality and procedure in DSM extraction. Specific focus is placed on SfM and its vertical accuracy performance when compared to the conventional techniques SfM was tested against in this study.

2.1 STRUCTURE FROM MOTION

Structure from Motion (SfM) is a technique for estimating the 3D geometry of object features from a collection of 2D imagery. The recovery of this structural information is solved typically by feature matching. By matching features a point cloud is produced which can be further enhanced by increasing the detail in a densification process. For practical application in geographical science, a surface is generated from the collection of derived points to produce an elevation raster.

2.1.1 Feature matching

SfM gained popularity due to its ability to match images from a collection of unordered images with no prior knowledge of the camera(s) needed. The identification of image features is used to match images using Scale-Invariant key points or Features which gave rise to the Transform algorithm (SIFT), a product of the work done by David Lowe (Lowe 2004).

Arguably the origins of SfM stem from multiple sources prior to SIFT which form key parts of the SfM workflow. Conceptually, the coplanarity conditions used by SIFT estimate the spatial relationship among images was initially applied in the 1960's for triangulation of aerial imagery (Thomson 1965) and bundle adjustment was established later to determine the relationship between the image and the identified objects (Brown 1971). There are multiple image matching operators like SIFT in existence such as Gradient Location and Orientation Histogram (GLOH), Speeded Up Robust Features (SURF), Least Discriminant Analysis (referred to as LDAHash) and more (Strecha et al. 2011; Harwin & Lucieer 2012) used in SfM. In a study comparing these algorithms, SIFT was found to be the most robust with large image variation in terms of untextured surfaces, occlusions, illumination changes and acquisition geometry (Harwin & Lucieer 2012). An emphasis is placed on SIFT as it is an integral part of the SfM processing schema used by Pix4D mapper (Tsay & Lee 2012) which is the software package used to perform this study.

The four major stages of the SIFT computation are: (i) scale-space extrema detection, (ii) key point localization, (iii) orientation assignment, and (iv) key point descriptor (Lowe 2004). In brief, the SIFT transform searches for image scales and locations using the difference-of-Gaussian function to determine key points of interest in the images (Lowe 2004). These key points are tallied to create candidate image locations and used to compute a model that matches points based on their stability. An orientation is assigned to each location based on local gradient direction, providing invariance to the transform as the scale, orientation, and location are processed for matching (Mikolajczyk & Schmid 2002). Lingua et al. (2009) noted that the SIFT algorithm is highly effective compared with other photogrammetric feature matching techniques.

Commercial applications with SfM can take on "black box" characteristics with proprietary blend and alterations to existing algorithms which is both an advantage and disadvantage (Machletti et al. 2015). The advantage is that expert knowledge is not needed to generate structural information but conversely the user has less control over the quality of the output and error may be harder to identify.

2.1.2 Key points to 3D point cloud

Matching key point descriptors between image pairs is achieved with the use of Arya et al.'s (1998) method for approximating the nearest neighbour using kd-trees while maintaining Lowe's ratio test (Lowe 2004). By applying Random Sample Consensus (RANSAC) (Bolles & Fischler 1987), the method can then estimate the F-matrix (Fundamental matrix) which is used to describe the epipolar geometry. The F-Matrix represents the relationship of corresponding points within stereo images. The feature points that exist within two images, one point is projected onto the other image constrained to a line, allow for the detection of incorrect correspondence between points. The matrix is optimised by performing a batch non-linear minimization (Bundle Adjustment), typically by using the Levenberg-Marquardt algorithm to solve the non-linear least squares problem to estimate the trajectory or 'motion' of the camera (Bolles & Fischler 1987; Snavely et al. 2008, Feetham et al. 2014). RANSAC will attempt to find the correspondence between candidate points even if they are incorrectly estimated, often resulting in outliers. The process will keep fixing/fitting the model (F-matrix) until the points connected are consistent; resulting in inliers with geometrically consistent matches (Bolles & Fischler 1987). The image pair matches are then organised into tracks (linkages between points across multiple images) by corresponding key

points across multiple images producing a connectivity graph (Furukawa & Ponce 2007; Furukawa et al. 2010).

The algorithm recovers camera parameters such as focal length and rotation for each image which is a particularly powerful attribute in instances where data was collected from more than one camera such as crowdsourced data (Cornelis et al. 2008; Snavely et al. 2008). The minimization associated with the camera parameters' reprojection results in a non-linear least squares problem that is solved using bundle adjustment. Bundle adjustment is a well-understood non-linear objective function with no closed form solution using non-linear least squares algorithms to minimize the objective function (Chen et al. 2012). Bundle adjustment is performed on best-fit image pairs and then repetitively with each image added along the track. This process is known to be time consuming with no real time solution to solve the minimisation restriction (Chen et al. 2012), but workaround techniques do exist to circumvent this issue (Newcombe & Davidson 2010; Newcombe et al. 2011). Within RANSAC, the camera parameters are optimised using a Direct Linear Transform (DLT) (Bolles & Fischler 1987). Camera triangulation is subsequently performed to calculate image position and the extent to which it captures the scene. Once all significant images are added, a final bundle adjustment is performed across all images (Snavely et al. 2008).

The success of matching descriptors is based on good coverage, texture and overlap. Agreement between images and completeness of scene yield better results. It is suggested that a minimum of 50% overlap between images is sufficient in identifying key points for matching stereo pairs and any feature of interest should appear in at least three images (Frankl et al. 2015). More images are recommended with greater overlap to ensure the successful generation of the model with a higher level of information (Photosynth 2005; Tavani et al. 2014). This does, however, increase processing time.

2.1.3 Point cloud densification

To increase the detail of the point cloud, a Multi-view Stereo (MVS) process is used to find more points with the recovered camera parameters and image orientation. MVS is a conceptually natural extension of stereo photogrammetry. Instead of using only two-view stereo algorithms, MVS uses more images in-between pairs to increase robustness of point estimation. With this extension, MVS

is also designed to cater for more varying viewpoints and image set sizes even in the order of millions (Tsai 1983, Okutomi & Kanade 1993).

The implememntation of Clustering view for Multi-View Stereo (CMVS) and Patch-based Multi-view Stereo (PMVS) process sparse point clouds into more meaningful renderings (Furukawa & Ponce 2007; Furukawa et al. 2010). The former increases the density of the points but clusters them into subsets of manageable size with the latter used to reconstruct a 3D representation of the feature (Matiukas & Miniotas 2011). PMVS only models rigid/static objects meaning that artefacts that appear as a result of natural interference (walking pedestrians, moving cars etc.) are for the most part not computed into the model (Furukawa & Ponce 2006). This is useful in its application for crowd sourced data and in street level photography as perfect conditions cannot always be achieved (Agarwal et al. 2011). Multiple MVS densification algorithms exist and have been evaluated (Seits et al. 2006). The method proposed by Furukawa & Ponce (2006) is one of the most widely used, typically in conjunction with open source software (Furukawa & Ponce 2006; Favalli et al. 2012; Westoby et al. 2012; Tavani et al. 2014) producing impressive results with no substantial cleaning required (Stal et al. 2012).

2.1.4 3D meshing

Multi-view reconstruction has been found to be comparable to laser scanners with a certain level of accuracy and completeness (Westoby et al. 2012). Producing a complete and complex scene is still, however, an open problem (Jancosek & Pajdla 2011). There are multiple methods of approach when computing a complete 3D representation or mesh of a large and complex object derived from point clouds. CMPMVS (Centre for Machine Precision's Multi-view Stereo) is an augmented Labatut et al. (2009) method that has shown superiority when compared to plane-sweeping, stereo or growing-based models (Jancosek & Pajdla 2011). Part of MVS's success is owed to the observation that derivative point clouds result in better recovery of weakly supported surfaces (reflective or semi-transparent objects) in a natural environment. A feature that may have high specular reflection, degrees of transparency or lack of texture will typically result in a very poor point cloud rendition and when reconstructing the scene, anomalous behaviour can arise when generating a surface mesh (Jancosek & Pajdla 2011). Anomalies typically arise when reconstructing plain walls, cars, windows etc. in an urban scene. The adaption of Labatut et al. (2009) method achieves its success from Delaunays tetrahedralization of points to identify point

occlusion within images and agreement among points that are weakly supported, deriving better surface tessellations (Jancosek & Pajdla 2011).

To take advantage of parallel, multithreaded and many-core processors the Compute Unified Device Architecture (CUDA) was developed by NVidia for high-definition 3D graphics using a programmable Graphic Processor Unit (GPU) (CUDA 2013). This provides enhanced computational power with a high memory bandwidth rather than data caching typical in CPUs (Central Processing Unit). As a result some software providers, both proprietary and open source (Designing Reality, CMPMVS, Point Cloud Library, etc.), have made use of, and are programmed with, CUDA resulting in the necessity for enabled graphics cards placing hardware restrictions on open source SfM processing (CUDA 2013). CUDA can be emulated with non-NVidia hardware but the recommended Compute Capability is rated at 2.0 for SfM which the emulators (e.g. CUDA waste) do not achieve effectively (CUDA 2013). This places hardware restrictions for these software applications in systems without compatible hardware. This serves as a motivation to use proprietary software that does not require, but may still utilize, this technology. A study by Hirschmuller et al. (2012) found that leveraging GPU-based processing decreased computational time and also allowed for real time depth map production.

2.2 STEREOPHOTOGRAMMETRY

One of the oldest, semi-automated and computer aided techniques still employed today is stereophotogrammetry and is well established in the field of GIS. Stereophotogrammetry is the technique of estimating geometric structure from a collection of photographic image pairs.

The field of photogrammetry is believed to have been created in 1851 by Colonel A. Laussedat of the French Battalion of Engineers in an attempt to generate topographic maps without large amounts of manual labour (Albertz 2007). This was realised with the invention of photography and mathematical perspective techniques from which he was able to generate the first proof with a map of Paris. Today we are able to produce surface models of large areas with advanced algorithms and digital cameras in conjunction with modern computers with this technique. Information to ensure the successful generation of these models requires multiple inputs including camera model, interior and exterior orientation, calibration information, GPS location and tie points and ground control points (GCPs) (Geomatica Orthoengine 2003).

The process of deriving surface models can be summarized into five main processes: Image Capture (interior orientation), Aerial Triangulation (Exterior orientation), Image Matching, Mass point production and Interpolation.

2.2.1 Image capture

Surface models derived from aerial imagery using traditional stereophotogrammetry are not without limitations both as a function of the sensor and the platform itself. Imagery from camera sensors is subject to distortion due to the curvature of the lens, focal length and various perspective effects (Geomatica Orthoengine 2003). This camera calibration information is typically provided by the vendor or the manufacturer of the camera used to collect the imagery. The calibration information is used to determine the interior orientation of the camera (perspective centre, focal length, radial and decentring distortion of lens assembly, image resolution etc.) which describe the metric characteristics of the camera.

Aerial imagery taken from piloted aircraft may be prone to radiometric and geometric error. Radiometric error is present from the physical sensor and in instances of unfavourable atmospheric conditions (i.e. haze) which is more typical of satellite platforms, particularly for spectrally based processing (Honkavaara 2009). Variable topography can induce relief distortion within an image associated with camera geometry in imagery taken from aerial platforms. This distortion is also influenced by the external geometry and focal length of the camera. The effect is reduced with imagery taken at a higher flying height (assuming an identical sensor) but is a function of the extremity of ground relief variability (Dietrich 2014).

To extract surface models there is a recommended overlap criterion of 60% along the main strip of images with a side overlap 25-30% (LPS Project Manager User Guide 2009). If not met it could result in the failure of image matching or lower quality surface reconstruction. Increasing overlap with increasing numbers of tie points can contribute to better error detection. The overlapping imagery remains in an image coordinate system with intrinsic internal calibration error and misrepresentation of topography. The interior calibration error refers to the inaccuracies of metric characteristics related to the imager such as focal length, image resolution etc. which is mainly used in photogrammetry to define the perspective centre and the radial distortion curve. The correction of this error is solved in utilizing GCPs, tie points and existing elevation information (Ghosh 2005; LPS Project Manager User Guide 2009).

2.2.2 Aerial Triangulation

GCPs are collected throughout each image with tie points contained within the overlap between image pairs. GCPs establish the relationship between the image and the ground by providing ground coordinates for points/features in the image. This coordinate information can be collected from theodolite surveys, total survey stations, GPS and orthorectified imagery with a DEM. Tie points link images together which can be automated if the original images contain ground coordinate information (Ghosh 2005).

This coordinate information is used to compute the exterior orientation of the imagery and establish the relationship between the earth's surface and the angular/rotational elements affecting the camera at the time of capture. This external orientation establishes the relationship between the image and object space. This rotational information can be captured in conjunction with the imagery with an on-board GPS. With bundle adjustment, blocks of images can be computed simultaneously using least squares adjustment to estimate unknown parameters and minimise the distribution of error. Bundle adjustment can be used for large areas with fewer available GCPs to compute the exterior orientation (Ghosh 2005).

2.2.3 Image Matching

With known internal and external orientation information, image matching can take place in the overlapping regions of the image dataset. This process seeks to match corresponding points extracted by the image matching method used. Common image matching techniques used in stereophotogrammetry are:

- Area Based Matching
    - Cross-Correlation
    - Fourier Methods
    - Local Methods
- Relation Based Matching
    - Resampling with Gaussian Filter
- Feature Based Matching
    - Moravec
    - Dreschler
    - Förstner

(LPS Project Manager User Guide 2009, Joglekar & Gedam 2012)

Image matching algorithms are classified into three principal groups: Area Based Matching, Relational Based Matching and Feature Based Matching (Heipke 1996).

Area- or signal-based matching determines correspondence based on similarity of grey values (Joglekar & Gedam 2012). As the name suggests it samples pixels from a search window based on neighbouring values. Principally a level of increasing similarity will exist between two views of the same object with courser scale (Joglekar & Gedam 2012). As such it is a correlation-based method because of the nature of determining correspondence with either a window or, in some cases, entire imagery as a reference. A regular window may not suit the registration of the imagery as any deformation by the transformation of geo-referenced imagery may not allow the coverage of the same corresponding part of the scene in the reference image and the sensed image. Correlation methods are sensitive to changes in intensity as a result of different viewpoints, noise and sensor type. Fourier methods are used particularly where varying conditions of the image scene are observed or where noise is prevalent in an image dataset to increase computational speed in preference to correlation methods (Vasuki 2014). This makes it a robust method for highly correlated and frequency dependent noise disturbances (Joglekar & Gedam 2012). Local methods are computationally fast and typically produce better results than Fourier methods with uniqueness constraints used for validation within the method in some versions (Muhlmann et al. 2002).

Relation or structure based matching is an automated technique for identifying image features based on their relationships between each other (Wang 1998). Using an image pyramid, it matches features from the lowest resolution to the highest, which have their own set of potential resampling methods (Shapiro & Haralick 1987, Hannah 1989).

Feature based matching determines correspondence between two image features and matches them as points (Forstner 1986). Like SIFT, featured based methods are constrained by the contrast/texture of objects in the observed scene making homogeneity or low contrast a limiting factor when tracking points across imagery (Lowe 2004).

2.2.4 Mass point production and interpolation

Digital surface models derived from conventional photogrammetric methods have a high positional accuracy but typically contain a lower level of structural detail of objects than that of LiDAR. The detail quality of features, such as vegetation and buildings, are attributed to methods of point cloud generation (Mueller 2013).

Mass points are generated with image matching methods based on image correlation (LPS Project Manager User Guide 2009). These mass points are typically collected as part of an automated tie point collection process and increase the level of detail. If enough tie points are identified a rough DEM can be generated. For more accurate representations of the observed surface, an interpolation stage is further employed to estimate more points and create the surface raster.

The estimation of the point coordinate is typically informed by predicting surface statistics (geostatistical) or measurement of the extent of similarity to surrounding points/cells (deterministic). Inverse Distance Weighted (IDW), Natural Neighbour (NN), Spline etc. are examples of deterministic methods that generate new points and assign values based on the surrounding points and in some cases specify parameters that stipulate the level of smoothing of the surface at that point. Geostatistical methods, including Trend and Kriging interpolators, compute the statistical relationship of the points producing a prediction surface with measurable certainty (ESRI 2012).

These photogrammetric surface models generated after interpolation are largely 2.5D, only allowing one measurement of elevation per unit area. Practically this makes it impossible to capture detail on flat vertical features. For example, walls of buildings in a built environment cannot be represented but only the ground and rooftops can be captured. If a surface is estimated from two points, one on the roof and one on the ground, to try and recover the change in surface height, interpolators will typically compute a weighted curve where the true representation of the surface is a sharp discontinuity of a near ninety degree angle (ESRI 2012).

Object proximity (relative to the spatial resolution) can also determine object separation and detail as objects can occlude detail and may generalise into one singular object instead of a collection of objects. This error is a factor of the sensitivity of the image matching technique, surface recovery algorithm and spatial resolution of the imagery.

2.2.5 SfM as an extension of Traditional Stereophotogrammetry

SfM and traditional photogrammetry are both techniques that recover structural information from imagery and as such utilise common photogrammetric methods as part of their processing schema.

Traditional photogrammetric techniques (TPTs) require calibration information and coordinate information as part of its initial processing from the user whereas SfM does not require initial camera information at all and coordinate information can be introduced either pre- or

postproduction of the orthorectified image mosaic/DEM. GCP and tie point information is needed to orientate and link images for TPT. This is not to say SfM does not require this information but rather that this information is recovered by the workflow introducing a level of automation from the different use of feature based matching.

Feature based matching is native to both SfM and stereophotogrammetry. There may be speculation as to the difference between current commercial photogrammetry and SfM as it can take on 'black box' characteristics when adopting proprietary processing workflows (Remondino et al. 2014) while maintaining these common algorithms. Although some software packages have begun to introduce SIFT for feature matching, the algorithm is still not yet commonplace for point matching. With feature identification, SfM and TPT are still exposed to the same environmental factors.

The success of matching descriptors in SfM is based on good coverage and texture as is stereophotogrammetry. In a comparative study by Lingua et al. (2009), SIFT was evaluated using terrestrial and UAV data with various matching techniques used in photogrammetry – the Förstner operator and Least Square Matching. The study found that SIFT was able to homogenise points in image pairs with high accuracies with image rotation ranging 20-30 degrees even with large geometric and photogrammetric distortion. It was also found that the number of points extracted by SIFT was higher than that of traditional techniques, which is advantageous for tie-points to correctly determine orientation for automated processing. It may have issues with feature distribution across a scene which is directly related to the texture available in the image data. This can be minimised with the utilisation of Adaptive SIFT which optimises its parameters to suit the image data (Lingua et al. 2009).

2.3 LIDAR

LiDAR (Light Detection and Ranging) is a laser based technique used for recovering accurate structures of objects, features and surface topography. DSM creation from LiDAR data is a well-established practice with a multitude of applications (Moreira et al. 2013). This technique, despite using parts of the Electro-magnetic Spectrum (EMS), does not rely on imagery to recover structure. It is an active sensor that emits light pulses and records the interaction between the light pulses of the observed surface to recover this structural information in the form of point clouds.

2.3.1 LiDAR sensors

LiDAR emits a laser pulse which interacts with an object/feature eliciting a change in the wave profile of the energy pulse which intrinsically contains range information. Modern devices typically observe measurable information of numerous surfaces in the form of multiple returns contained within the wave form. The pulses are comprised of high energy particle (photons) emissions focused towards the earth surface with a Pulse Repetition Frequency (PRF) (Mallet & Bretar 2009). Photons are reflected back to the photodetector which records the intensities of the wave profile per point which, in turn, is translated into surface and terrain detail. For airborne topographic measurements a narrow wavelength of the electromagnetic spectrum (EMS) is generally measured in the near infrared range at 1064nm in single wavelength systems (Mallet & Bretar 2009; Wang et al. 2014).

LiDAR pulses are recorded in two ways either discrete return systems, or full-waveform system. Discrete return systems emit pulses and collect the number of individual returns present in the wave-form. This pulse may initially observe an object and part of it will return while some will pass through the object and return when it observes either other parts of the object (i.e. tree branches) or other objects (Jensen 2007). Full-waveform systems, however, do not record pulses as returns but rather the amount of energy returned measured over a period of time by the sensor at equal intervals, this information is referred to as intensity.

Flood in 2002 identified five degrees of processed LiDAR data provided by vendor's determined by the needs of the user's applications:

1. Basic/All-points
2. Low fidelity/First-pass
3. High fidelity/Cleaned
4. Feature layers
5. Fused

The initial three degrees of processing have different levels of filtering but contain no feature identification information. Levels four and five have this extracted feature information, as determined by the vendor, but typically takes longer to derive thus information and may cost more to produce (Cheuk & Yaun 2009). Similar to SfM, the processing of raw from specific LiDAR sensors may be proprietary and unstandardized making the understanding of processing steps

related to the instruments critical to limit potential introduction of error and application (Deems & Painter 2006).

2.3.2 Topographic LIDAR and interpolation

The information that LiDAR systems provide is in the form of point clouds which don't intrinsically have real world coordinates associated with them. This additional spatial information is derived with a GPS and an Internal Measurement Unit (IMU) which records positional and orientation information respectively. The IMU specifically records information regarding the aircraft's orientation and by extension that of the LiDAR system (Mallet & Bretar 2009). This information is comprised of roll, pitch, yaw and altitude which is collected as perfect nadir flight is impossible. Inconsistent aircraft orientation is due to a range of factors including human error, mechanical performance and atmospheric effects.

Height information is derived in most LiDAR systems by utilising a Time of Flight (TOF) process to determine the distance between the sensor and the laser-illuminated object in a post processing step (Mueller 2014). Each LiDAR system has a recommended maximum altitude (or range) and view angle specific to the unit. The closer to the ground the LiDAR flies, the more points are captured per square meter. The laser's swath may have a marginal effect on sample distribution with a gradual increase in sample spacing as points are taken further and further away from the aircraft assuming a flat surface. This increase in sample spacing effect is dependent of the individual LiDAR unit used.

The georeferenced point clouds need to be filtered and interpolated into a surface raster of the elevation values. This filtering process separates the returns observed by the LiDAR sensor to derive the information layer required. The first return(s) contain information of surface and objects contained on the surface whereas the last return contains the terrain without the surface object. This is not true in all cases as hard objects (buildings, bridges, rock formations etc.) will have a similar response to the first return as they are impermeable by the laser beam. This filtered information is interpolated to increase the detail by estimating new points in areas of no information. Similarly to stereophotogrammetry a variety of purpose specific methods are available.

2.3.3 SfM and LiDAR

Data derived from SfM is not from a laser beam but a series of images relying on overlap to recreate feature point clouds. SfM has no returns other than the feature it recreates and will thus have no directly derivable ancillary data about the land cover resulting in point clouds existing in an arbitrary coordinate system. This necessitates georeferencing for it to have real world coordinates which may come in the form of geo-tagged imagery or GPS location data. The intrinsic difference is that SfM relies on texture within an illuminated image pair to identify features whereas LiDAR is dependent on the density of point capture and the features being captured. Terrestrial Laser Scanners (TLS) have been found to be comparable to the spatial results of SfM with ground-based imagery (Westoby et al. 2012). Airborne Laser Scanners (ALS) have similarly been found to be comparable to SfM from UAV imagery; both in point cloud density and with accuracy in the centimeter range (Fonstad et al. 2013). The focus of this paper was to test accuracy in multiple environments to see if it held true across differing terrains in study sites.

2.4 DSM CREATION AND ERROR PROPAGATION

DSMs are typically created from points or contours derived from raw data sources. Any surface measurement or estimation technique has an intrinsic margin of error associated with the instrument or algorithms used in the recovery of elevation information (Hengl & Reuter 2008). The identification of the possible sources of these errors, and their mitigation, are important to the quality of the final DSMs (Oksanen 2006).

In a typical multistage process error propagates and may increase with every stage of processing. This will amplify the error in the final output and increase the uncertainty in the accuracy of the derived information.

Figure 2.1 illustrates the uncertainty and error that may propagate throughout a dataset with each processing step from the point of capture to the application and interpretation of the data. This depiction shows that every operation performed on a dataset may have an additive effect on error and uncertainty. Uncertainty in derived datasets increases with each processing step when the measured propagative effect of error is unknown.

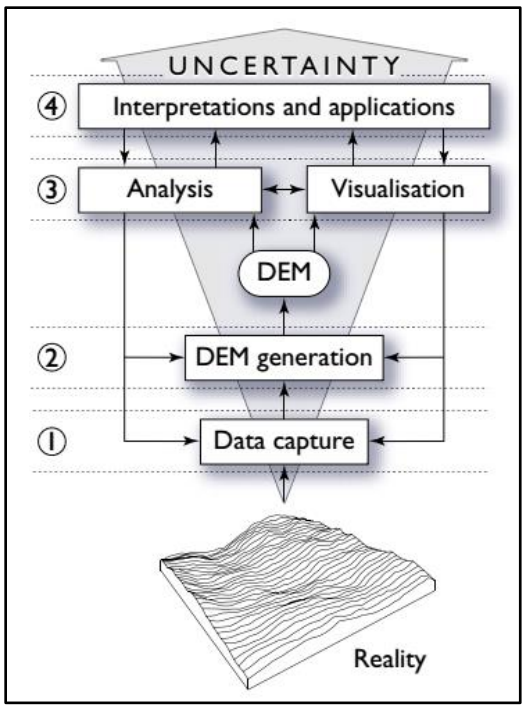


Figure 2.1 shows uncertainty within data, depicted as an increasing function of data manipulation, as an illustration of error propagation (Oksanen 2006)

In the model, illustrated in Figure 2.1, a feedback loop demonstrates various steps to identify and limit the propagative effect of error using additional information as it is acquired from derived datasets.

The uncertainty of the error, amount or nature, may be introduced into the decision making in one of four phases and may increase with progression. In phase 1 random error and human error can occur in the data capture and acquisition. Examples of this would be measurement error in remote sensing equipment, GPS and inertial navigation systems, and geometric transformations. In phase 2 the method of approximation and the level of detail used to represent the DSM, such as cell size, point density etc., can contribute to error. In phase 3 uncertainty can be introduced in the analysis originating from the algorithms selected and the assumptions made (Zhou & Liu 2004). Lastly, in phase 4, uncertainty may be caused by misinterpretations and incorrect application by the user. By using this multi-phase approach to identify sources of error, and by extension their propagation, it is possible to inspect the DSM generation process to find where error is introduced from data capture to the final product and apply appropriate corrections or adjustments.

2.4.1 Sources of error in raw data

Errors that occur in primary data may propagate with each derived dataset increasing the magnitude of the errors and the uncertainty and reducing confidence in the resultant DSM. The sources of error in the raw data vary according to the collection techniques used.

Optical methods rely on contrast which is used to recover structural information about the observed surface (Campbell 2006). The resolution will affect the quality of the raw imagery and may introduce uncertainty within the data (Oksanen 2006). In imaging techniques, such as SfM and TPT, the spatial resolution of rasterised datasets can result in the mixing of spectra, referred to as 'mixed pixels'. This is the averaging of several spectral classes within a pixel space, and is present where spectrally dissimilar features meet and the spatial resolution is not fine enough to clearly define these boundaries. The coarser the spatial resolution, the higher the proportional number of these mixed pixels (Campbell 2006). Spectral resolution describes the ability of a sensor to define wavelength intervals. The narrower the wavelength range for a particular channel or band, the higher the spectral resolution (Campbell 2006). Commercial cameras typically capture RGB where the wavelength range may be quite large, capturing substantial portions of the visible spectrum. These wavelengths are generally averaged to a single spectral value and as a result the spectral contrast of more discrete features and overall brightness may be lost (Bemis et al. 2014). Brightness values define contrast within a band and represent the magnitude of the electromagnetic energy measured. Spatial resolution can also affect the radiometric resolution as the magnitude of energy values are the collective average of the spectral response. This is based on the spatial occupancy of each pixel relative to the ground influencing the value of this magnitude in the value range (i.e. 8bits = 0-255 value range). This information refers to the amount of information within each pixel potentially available from which detail can be extracted in the form of a digitally stored number. The raw imagery may only capture a smaller range of numbers or record mixed brightness values which would supply the surface generation algorithms less information to recover elevation information compared to that of imagery with a higher number of bits.

For DEM creation imaging techniques rely on feature contrast. Errors that result in degraded or low contrast will reduce the sensitivity of the information recovery process which leads to miscalculations and smoothing or over-generalisation of features (Bemis et al. 2014). This is a significant problem in horizontal accuracy. In terms of vertical accuracy this source of error may not be as detrimental or impactful depending on the scale and detail required by the project. For

example, the ability to reconstruct pebbles in a riverbed is nonessential for most national level DEM processes and can rather be seen as project specific requirements.

LiDAR may produce spurious results derived from photon counting statistics with regard to calibration coefficients and background corrections. Sources of error in raw LiDAR may arise from positional errors which may not be attributed to the sensor itself but rather the GPS leading to incorrect spatial information given to the measurement unit (Hodgson & Bresnahan 2004). This spatial information may be tainted, not necessarily by inaccuracies of the positioning units, but also when the LiDAR sensor is not perfectly aligned with the platform. A simple example of this is vibration. Small alignment errors (off focus, off axis or misalignment in system pointing direction will lead to part of the backscattered light being lost in the optics, detector, or atmosphere. Error, that may be present in the laser system itself, may arise from false atmospheric corrections (such as those from inaccurate photon counting statistics), or range error due to inaccurate time measurements and ambiguities in the target surface, all of which produce what is known as a 'range walk' (Hengl & Reuter 2008). The object or surface complexity may have a higher probability of introducing uncertainty as the sensing of varying reflective properties and geometry (specular reflectance) of the surface features when capturing coordinate information may not accurately be captured. This can increase possible spurious data points resulting from multipath reflections. The LiDAR sensor scanning angle can also lead to less information being captured towards the edges of a sensed scene, providing less information to the periphery. (Hodgson & Bresnahan 2004).

Although these errors in primary data do not propagate intrinsically, without removal they may propagate with every further stage of processing. Any derived data will carry with it these intrinsic errors which may worsen with every operation making the dataset more inaccurate, error more pronounced and decrease the confidence in the data.

2.4.2 Sources of error in derived data

The quality associated with DEM generation will vary with the input data as well as the chosen method of processing used (Hengl & Reuter 2008, ESRI 2012). The derivation of elevation is typically generated as a point dataset which is used to derive the DEM. Depending on the method used to generate these points, a level of filtering may be required before the actual surface is created. The point generation technique, filtering and gridding (rasterising) may all contribute to errors and uncertainty within the data.

In the case of SfM, a large amount of elevation information is generated as a sparse point cloud and used to further re-assess the imagery and to find more points within a scene from the imagery in a densification process (Furukawa & Ponce 2007). SfM points are estimations of point data of a specific feature taken from overlapping images and multiple points can exist at a specific location with varying elevations. Unlike LiDAR, these elevation points are not present in a waveform so they cannot be as readily separated out. Only one value can be used as a single elevation data per unit area when gridding. To achieve this a filtering process is added to determine a single height value per DEM pixel. Depending on the algorithm used, an average, minimum, maximum or trends can be used to determine the final elevation value. In doing so, an estimated elevation value has to undergo another degree of estimation. This is often valuable for removing artefacts and erroneous points but also introduces uncertainty as the filtering process removes raw data in favour of a single estimated value, which may not be representative of the observed surface. SfM software typically utilizes a Triangulated Irregular Network (TIN) instead of an interpolator to generate surfaces without generating more points. These networks are then converted into a cell grid, or raster, introducing possible averaging errors depending on the recovered detail and spatial resolution of the cells required (ESRI 2012).

Elevation information is recovered by TPT by determining epipolar geometry to create point heights. Some of these points may have to undergo a filtering process to increase the quality of the elevation dataset. These filters are used to remove noise, erode holes in the data or smooth the surface points. TPT typically uses a pixel-based window (i.e. 5x5 pixel window) to correct targeted elevation values using, for example, median or Gaussian filters (Geomatica Orthoengine 2003). This filtering process increases the quality of the derived surface but may remove features that are actually present on the surface and may erroneously overgeneralise or smooth objects.

Elevation point data is generally processed into 2.5D grid format which allows users to have an elevation value across an entire scene without gaps. Processes typically create a grid by using the existing surface values to predict elevation of cells at locations that lack a sample point based on a form of spatial autocorrelation or trend. Depending on the gridding method, or interpolator, these processes may change the original value based on the mathematical procedure used. There are two categories of interpolators: deterministic and geostatistical or stochastic. Deterministic interpolators are based on measured points and the extent of similarity between cells whereas geostatistical interpolators are based on statistics to predict the surface model. Cells that are

unmeasured rely on statistics or measured points to generate a value which introduces a degree of uncertainty based on estimated values that may not necessarily be representative of the real world surface. Geostatistical interpolators generally provide a measure of confidence in the generated surface whereas deterministic interpolators typically do not. This does not mean that geostatistical interpolators are superior as there are benefits to each interpolator type (ESRI 2012).

There also exists the possibility, given the spatially rigid nature of raster cells, that point elevation values may be incorporated into the same cell further re-estimating the elevation at that location.

These predicted values and re-estimated elevations, along with intrinsic error in the primary data, propagate and increase the level of uncertainty within the DSM as represented in (Figure [todo]) and can be present in SfM, TPT and LiDAR (Oksanen 2006).

Lastly, in comparative studies, temporal effects may affect the analysis of DSM comparatives and derivatives (Florinsky 1998, ESRI 2012). Natural and human activity can change the terrain of a given area and can impact comparative results depending on the degree of change. This effect is dependent on the observed area and how stable surface features are. For example, imagery in summer versus winter may have a marked difference in surface profiles in vegetated areas but may remain static on stationary dunes or cityscapes.

2.4.3 Post DSM data cleaning and error removal

Even the most sophisticated DEM generating techniques can still contain errors which may need to be removed. Their removal can limit the propagate effect of error on datasets that derive information from the DSM. However, any error not addressed prior to the cleaning of this elevation data may still be present depending on the effectiveness of removal. There are three main errors that may be present within a DEM dataset: gross errors, systematic errors and random errors or noise (Hengl & Reuter 2008).

Gross errors or artifacts are typically easy to detect visually usually presenting themselves as missing or grossly exaggerated values contained within the DEM. Systematic errors unlike artefacts, are not clearly erroneous and visual inspection may fail to identify them. In the case of DEM's produced by TPT, shaded, (low reflectance due to shadows) outer, edges of the image pairs may produce higher levels of inaccuracy. Random errors or noise may result in spurious information that can present itself in multiple forms and identification may also not be picked up

visually. Similar processes in handling these errors can be applied including windowed averaging methods.

2.5 COMPARITIVE CASE STUDIES

Despite a lack of research into accuracies associated with SfM in GIS, there are comparative studies in other limited and purpose specific research. These can reasonably be expected to provide accuracy ranges in differing environments and across differing techniques applicable to this study. Within the current research, factors affecting the reconstruction quality of SfM surface models were identified. Any technique that measures or estimates a surface will have error, the estimation of acceptable error or 'error budget' can be determined by pervious findings (Hodgson & Bresnahan 2004). These comparatives inform the expected findings and limitation of this study.

An investigation conducted by Mueller (2014) compared feature based matching from SfM, and stereophotogrammetry using aerial imagery, satellite imagery and ALS. The focus was on point clouds and their associated reconstruction accuracy. Using relative methods (no independent reference), it was found that both MVS and photogrammetric models could not compare to LiDAR in object reconstruction. This was illustrated in cases where SfM-MVS tended to round surfaces of rigid free-standing objects (i.e. buildings) and produced limited detail of vegetation. The assessment was largely qualitative and observational, drawing conclusions from cross-sections of the study areas. Conversely, a study by Westoby et al. (2012) had a conflicting conclusion in a comparative study of SfM and LiDAR in the field of geomorphology using terrestrial data acquisition. Terrestrial data is prone to 'dead ground' where surface mounds and vegetation can obstruct/occlude the collection of data behind the observed feature (Westoby et al. 2012). This results in pocket areas with no data being observed. Despite issues resulting from this dead ground issue and offset created by vegetation, SfM achieved decimeter-level vertical accuracy including sites with a complex topography concluding that SfM was indeed comparable to LiDAR. In a similar survey, Gómez-Gutiérrez et al. 2014 found centimeter-level accuracies while monitoring gully headcut erosion with average distances between the SfM and a reference point clouds to less than 2.5cm which is a marked increase in accuracy from Westoby et al. (2012). In urban environments, Dietrich (2014) study confirmed a high vertical accuracy to that of Gómez-Gutiérrez et al. (2014) when comparing SfM to TLS outputs.

Unmanned Aerial Vehicles (UAVs) have opened more affordable methods of aerial data acquisition and reduced error associated with dead ground using aerial imagery. The lower flying height of UAVs can provide higher spatial resolution to that of its piloted counterparts but with reduced field of view. This will in turn increase the imagery dataset size and subsequent processing time. Fonstad et al. (2013) compared vertical accuracy of SfM using low-altitude (10-70 m) UAV imagery with airborne LiDAR. Accuracy was measured by subtracting GPS points from both LiDAR- and SfM-derived DEMs. A relative accuracy was also computed by subtracting the ALS DEM from the SfM DEM. The results indicated accuracy of SfM in the centimeter range with a Root Mean Square Error (RMSE) of 0.07 m against GPS points and 0.51 m. When the aerial LiDAR was compared to the same reference, similar results were observed also concluding that SfM is a comparable technique.

Johnson et al. (2014), in a study of fault zone topography, performed a similar evaluation of SfM at around the same altitude (50 m). The reference data comprised a mixture of TLS and ALS surveys. Their distance computation showed that 90% of the LiDAR points deviated vertically from the closest SfM point by less than 0.41 m. The largest deviations were observed on the difference raster between LiDAR and SfM on steep slopes, the outer edges of the difference map, large bushes and active or recently active stream channels. With GCP optimisation, SfM accuracy was increased to less than 0.13 m. This illustrates that ground control optimisation can have a marked effect on accuracy.

Harwin and Lucieer (2012) tested the accuracy of georeferenced point cloud with a total station survey using a differential GPS to georeference SfM point clouds. Their results indicated that a georeferenced point cloud accurate to less than 0.40 m could be obtained from imagery acquired from a low flying UAV. In a similar study of glacial land forms, Tonkin et al. (2014) used 7761 total station survey points as a reference for SfM accuracy assessment to produce total vertical RMSE values of 0.25 m to 0.40 m. Errors as low as 0.20 m were observed for less densely vegetated areas of the DSM, representing a consistent accuracy range between both studies using similar methods. Vegetation appeared once again to be a constant source of error and uncertainty. A study on the modelling of the topography of shallow braided rivers by Javernick et al. (2014) also compared DEMs derived from UAV-based SfM and aerial LiDAR. Despite utilising vegetative noise reduction techniques and finding a 0.10 m vertical error in non-vegetated areas, the overall range was 0.13 m to 0.37 m. This was a particularly good result allowing for a 600-

800m altitude for imagery capture. It is comparable to the vertical accuracy measured by Hugenholtz et al. (2013) at much lower altitudes (200 m). Stal et al. (2012) produced accuracies above that of either Javernick et al. (2014) or Hugenholtz et al. (2013) at altitudes of 400 m with an RMSE value of 0.21 m. This further suggests the field of view (FOV) and the sensor play a larger role in reconstruction success and accuracy than flying height.

Structure from Motion is a function of the spatial resolution of features and its ability to resolve its structure. Image occlusion occurs when parts of an object or feature are not visible in the imagery due to oblique perspective or obstruction. Occlusion limits the ability of SfM and stereophotogrammetry to gather information about the terrain in a covered environment unlike LiDAR with multiple returns. This in turn impacts the vertical accuracy of SfM when compared to an independent reference.

The detail of the DSM is derived almost directly from the density of image points per area of an object. Leberl et al. (2010) concluded that the current literature placed the accuracy of UAV-based SfM around 0.20 m with the best cases producing models with an accuracy of 0.03m. Turner et al. (2012) performed an absolute spatial analysis from UAV data. Taking into account horizontal and vertical components, they found that SfM models deviated from the reference by less than 0.40m but with an overall result range of 0.103-1.247m between all the models.

Table 2.1 summarises the findings of the comparative case studies and gives some indication of expected accuracies for this study. When SfM was compared to total survey stations/GPS reference a vertical accuracy range of 0.07m-0.59m was demonstrated (Harwin & Lucleer 2012; Tonkin et al. 2014). This provides a baseline expectation for the absolute accuracy assessment. Where aerial laser scanner were used for validation a vertical accuracy range of 0.21m-0.41m has been shown for relative accuracy assessments (Stal et al. 2012; Johnson et al. 2014). Despite these ranges, some studies have reported accuracies as just less than a meter (Westoby et al. 2013). However, 90-94% of SfM derived data is expected to exist within a meter of the reference possibly due to the inclusion of outlying data points (Westoby et al. 2013; Gómez-Gutiérrez et al. 2014; Johnson et al. 2014).

Table 2.1 Summary of SfM studies illustrating vertical accuracy to independent references as validation

Study	Platform	Survey Altitude (m)	Validation	Vertical Difference
Westoby *et al.* (2013)	Terrestrial	Ground Level	TLS	94% of point values within ±1 m
Gómez-Gutiérrez *et al.* (2014)	Terrestrial	Ground Level	TLS	RMSE = 0.025-0.091m
Dietrich (2014)	Terrestrial	Ground Level	TLS	RMSE = 0.180-0.195m
Javernick *et al.* (2014)	Piloted Helicopter	600-800	GPS	RMSE = 0.13-0.37m
Hugenholtz *et al.* (2013)	Fixed-wing UAV	200	GPS	RMSE = 0.29m
Tonkin *et al.* (2014)	Multi-rotor UAV	117	Total Stations	RMSE = 0.200 - 0.588m
Harwin & Lucleer (2012)	UAV (not specified)	110-130	Total Station	RMSE = 0.25-0.40m
Fonstad *et al.* (2012)	Helikite	10-70	GPS	RMSE = 0.07m ALS mean difference 0.6m
Stal *et al.* (2012)	LiteMapper 5600 UAV	400	ALS	RMSE = 0.212m
Johnson *et al.* (2014)	Motorized Glider Helium Balloon	150-300 50-120	TLS+ALS	Unoptimised: 90% points values within 0.39- 0.41m Optimised: 90% points values within 0.13 - 0.32m

Expanded and adapted from Tonkin *et al.* (2014)

As Green et al. (2014: 181) stated "Until structure from motion can demonstrate reliable accuracy, and this can be calculated on a case by case basis, it is unlikely to be taken as seriously as a measurement tool". With more and more studies producing accuracy assessments, the confidence in the use of SfM might increase, although accuracy variability is still an issue. Collectively the current case studies suggest that a user can expect sub-meter vertical accuracies with adequately acquired imagery.

CHAPTER 3: METHODS

This chapter outlines the data required for the study and documents the methods of processing and analysis used to achieve the study objectives. For the purpose of clarification the methods are presented in five main processes: site selection, data acquisition, DSM generation, land cover classification and accuracy assessment. Figure 3.1 illustrates the workflow through the five main processes.

As the study is aimed at tests in different environments, site selection is used to identify a heterogeneous area encapsulating regions with different land use and land cover.

Data sourcing comprises the collection and capture of reference and study data used to derive DSMs for testing. Data collection is from satellite data, aerial imagery, LiDAR data and DEMs from various vendors available in the public domain.

With the collected and captured data, DSMs are generated with LiDAR, stereophotogrammetry and SfM using the relevant DSM processing schemas (Figure 3.1).

The accuracy of the DSMs is expressed in two ways: absolute and relative accuracy assessment. The absolute accuracy assessment is a quantitative measure of accuracy across all three DSM generation techniques. Relative accuracy assessment is a measure of accuracy to the best performing surface generation technique and to land cover/land use areas. Land cover is determined through a SPOT 5 classification process which has its own accuracy evaluation contained within this method section. This processing workflow can be followed in Figure 3.1 and in more detail throughout this chapter.

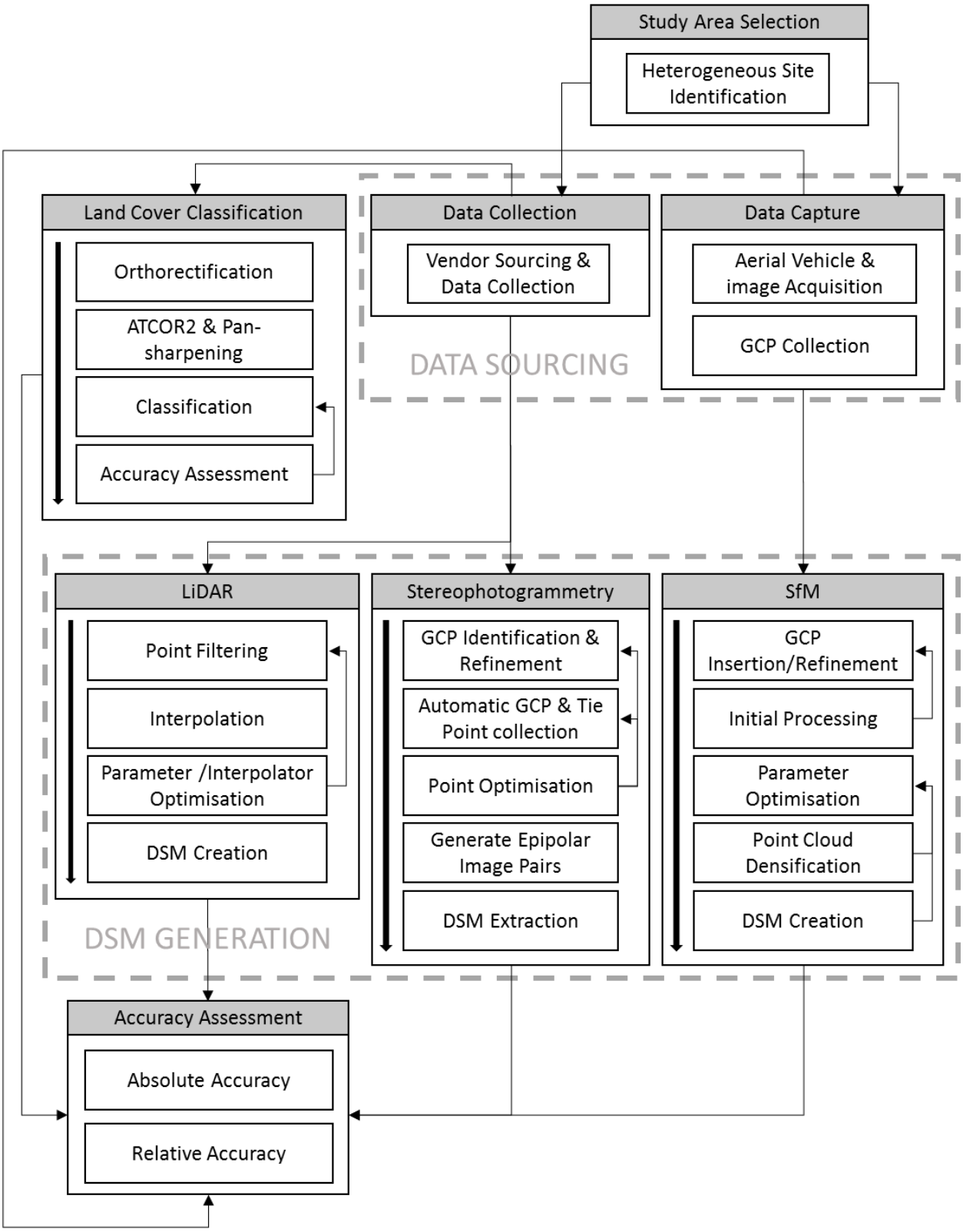


Figure 3.1 Workflow of the methods used in this study

3.1 STUDY AREA SELECTION AND EXTENT

The purpose of the study area selection was to assess LiDAR, SfM and traditional stereophotogrammetry for testing in multiple and diverse land use and land cover areas in the Western Cape, South Africa for this case study.

Considerations and limitations are imposed on the location as the study area has to fulfil several criteria including heterogeneity of land cover, accessibility, practical considerations and legal data collection. To ensure the quality of the study, an area with adequate heterogeneity was sought after to allow sufficient sampling across multiple land cover areas. The desired land covers to be sampled were varieties of built structures (e.g. tarred road, industrial buildings, residential homes, informal dwellings, etc.), both natural and planted vegetation (e.g. crops, wetlands, shrubs, trees) and natural and artificially exposed earth (e.g. dirt roads, bare soil, quarries, clearings). These sites had to be accessible by financially viable transportation options for manual data collection. The practical considerations were that manually captured data had to be contained by the data extents available by a vendor. This limited the area for collection to the Cape Town metropolitan area. The overlap criteria for SfM data for DSM construction are not met by vendors so manual capture was necessitated via aircraft. A functional requirement of using aircraft necessitates the presence of a local airport. Flying aircraft in the Cape Town metropolitan imposes flying restrictions due to the proximity to Cape Town International Airport which had to be taken into account for the legal flying of the study area. From this information, multiple candidate sites were identified and a single suitable site was chosen, approximately centred on the north east corner of Milnerton Rural, Western Cape. The Morning Star Airport was used as the terminal for the aircraft (Figure 3.2). Other legal considerations involved the acquisition of appropriate permissions for access to the private properties within the study area. Ethical and community permission was sought to access informal settlements. Despite the fact that informal dwellings are developed on public land, social permissions and notifications were needed to access land where people reside, catering for cultural sensitivities to non-inhabitants of the area. The ethical requirements were evaluated and approved by the University of Stellenbosch.

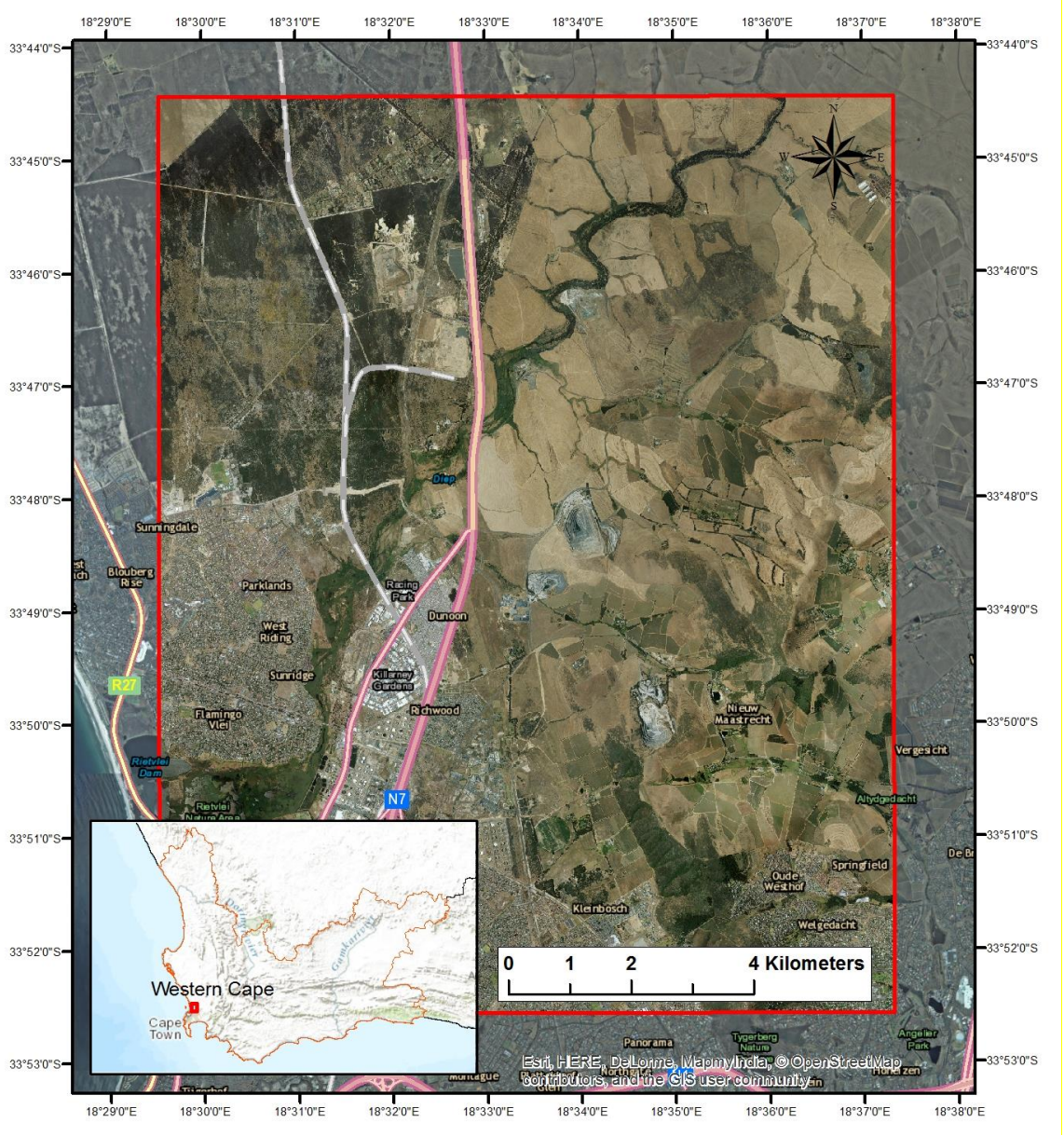


Figure 3.2 Selected study area of the greater Milnertion area, Western Cape, South Africa

The area of study was selected as the greater Millerton area situated north-east of the Cape Town CBD. The boundary spans the 1:10 000 orthophoto mapsheets 3318DC01, 3318DC02, 3318DC06, and 3318DC07 as defined by the CDNGI (Chief Directorate Nation Geospatial Services) and extends from 33°45'0.93"S to 33°51'0.57"S and 18°29'59.80"E to 18°36'0.01"E (Figure 3.2).

3.2 DATA ACQUISITION

The acquisition of data required for this study was separated into primary data capture and secondary data collection. The primary data consists of aerial imagery for input into SfM and differential GPS measurements for validation and georeferencing. The secondary data consists of existing commercial stereo aerial imagery as well as LiDAR point clouds. Secondary data also included satellite imagery from which a land cover map could be extracted. Pre-existing data, in the public domain, was sourced from various vendors depending on the nature of the data. SfM imagery and GCPs were captured manually as existing data was non-existent or insufficient as an independent reference.

Table 3.1 illustrates how the data was acquired with the associated vendor, resolution and date the data was created. The main contributor of the secondary data used for DSM generation was collected by the City of Cape Town's (CoCT) Geospatial Services department. Data needed for TPT triangulation process came from the Centre of Geographic Analysis (CGA) and CoCT Geospatial services.

Table 3.1 Data acquisition with respective vendors, resolution and capture dates

Data	Acquisition	Vendor	Spatial Resolution (m)	Date Created (mm/yyyy)
LiDAR	Secondary	City of Cape Town, Geospatial Services	0.35 - 0.62*	12/2013
Aerial Imagery	Secondary	City of Cape Town, Geospatial Services	0.12	11/2013
Reference Imagery	Secondary	City of Cape Town, Geospatial Services	0.08	12/2012
Satellite Imagery	Secondary	SANSA	Panchromatic: 2.50 (G,R, NIR): 10.00	12/2013
Reference DEM	Secondary	CGA	5.00	2012
SfM Imagery	Primary	Manual	0.09 - 0.15**	4/2014
Reference GCPs	Primary	Manual	0.10	8/2014

* Average scan line point spacing. Does not account for overlapping scan lines.

** Average ground sampling distance (images were captured with varying orientations and flying heights).

3.2.1 Secondary Data Collection & Processing Toolkits

LiDAR and aerial imagery was acquired from the CoCT Department of Geospatial Services. Both datasets were captured in the beginning of November 2013 for both data types and completed at the end of December 2013 (Table 3.1). The reference elevation model used for triangulation information in stereophotogrammetry with a 5 meter resolution product called SUDEM, acquired from the CGA (Van Niekerk 2012). The SPOT 5 multispectral satellite imagery was provided by the South African National Space Agency (SANSA) and used for the land cover classification, which is pertinent to the relative accuracy assessment.

Due to the variety of datasets multiple software packages, seen in Table 3.2, were needed for DSM production, DSM evaluation and ancillary data production. The choice of software package and the algorithms it uses can make a marked difference to the resultant output data. As LiDAR needs very little user interaction, ESRI's ArcMap 10.2.1 was sufficient to produce surface models. The traditional photogrammetric DSM production was performed in Geomatica's Orthoengine 2014. Pix4D mapper was used for SfM processing and surface production. Since multiple software packages exist for SfM processing (Pix4D mapper, Designing Reality, Agisoft, Smart 3D Capture etc.), the variety of proprietary and non-proprietary algorithms may induce variations of vertical accuracy between software packages. A combination of Geomatica's Orthoengine and Focus 2014 were used to orthorectify, atmospherically correct and pan-sharpen the satellite imagery. Trimble's eCognition with ArcGIS Arcmap 10.2.1 were used for the land cover classification. Graphpad Prism was used to analyse uncertainty as part of the relative accuracy assessment.

Table 3.2 Software packages utilised for data processing and analysis

Software	Function
Pix4D Mapper	Structure from Motion processing
PCI Geomatica 2014: Orthoengine	Traditional aerial photogrammetric processing & satellite image orthorectification
PCI Geomatica 2014: Focus	Satellite image post processing (Pan-sharpening and atmospheric correction)
Trimble Ecognition	Multiresolution segmentation & object based classification using SVM
Trimble's Pathfinder Office	Post processing differential GPS points
ArcGIS: Arcmap 10.2.1	Pixel based classification, LiDAR interpolation, raster calculation, zonal statistics & map creation
Graphpad Prism 6	Statistical processing of frequency distribution

3.2.2 Primary Data Capture

SfM data was collected manually as existing imagery did not satisfy the overlap criterion for effective SfM based processing (Snavely et al. 2008). UAV imagery capture was planned for SfM aerial data as it is well placed to solve the limitations presented by terrestrial imagery (e.g. dead ground). Unfortunately during UAV platform testing restrictions were adopted in South Africa and UAV usage was suspended pending the introduction of formal legislation. An alternate method was therefore selected to simulate UAV imagery. To ensure ethical and legal viability for this study, a gyrocopter was utilized to capture the SfM imagery with an appropriate camera equivalent (Figure 3.3). This change was considered an acceptable substitution for UAV image data.

Figure 3.3 Gyrocopter with pilot and Canon D600 camera operator

To simulate imagery taken from UAV platforms, Table 3.3 illustrates the flying height equivalents of a GoPro Hero 3 (lightweight camera for UAVs) and a Canon D600 (the camera used) to estimate spatial resolution prior to the flight. This equivalency was created using fixed focal lengths: 18mm for the Canon D600 and 16.5mm for the GoPro Hero 3. This approximation predicts that GoPro Hero 3 flying at 150m (492 ft) has the same estimated ground resolution as a Canon D600 flying at 510m (1673 ft).

The flying height was variable as height adjustments were made as a result of pilot input (Table 3.3). The average flying height was estimated at 364m (1194ft) above ground level with a range of 8m - 438m for the entire collected dataset. This correlates with a GoPro H3 flying at 100-120m (400ft) with an approximate spatial resolution of 8-10cm for the majority of the data. The current laws governing UAVs support 100-120m as an acceptable and legal flying height (South African Civil Aviation Authority 2015).

Table 3.3 Camera simulation in spatial resolution at different flying heights. GoPro: sensor size 6.25x4.69mm, resolution 4000x3000px. Canon D600: sensor size 22.30x14.90mm, resolution 5184x3456px.

Camera	Flying Height (m)	Flying Height (ft)	Res X (cm)	Res Y (cm)	Res Area (cm^2)
GoPro H3	150	492.13	13	11.5	149.50
	120	393.70	10.4	9.2	95.68
	100	328.08	8.7	7.7	66.99
	80	262.47	6.9	6.2	42.78
Canon D600	510	1673.23	12.2	12.2	148.84
	410	1345.14	9.8	9.8	96.04
	340	1115.48	8.1	8.1	65.61
	270	885.83	6.5	6.5	42.25

Features on the ground were captured as close to nadir as possible to limit surface occlusion from objects with the exception of residential environments where flying at these altitudes is restricted, necessitating the capture of purely oblique imagery (South African Civil Aviation Authority 2015). This remains adequate for SfM data capture in these areas.

Steep topographic features tend to have a poor reconstruction quality based on the camera's physical ability to generate sufficient coverage of the area in question (Westoby et al. 2012). This may also include dead ground (dips and crests) in images that are not captured. Effort was made to ensure that these limitations were minimised with multiple passes in specific areas with the inclusion of oblique imagery. This was particularly necessary in the quarries where illumination on one face of the pit was obstructed by the excavated walls. The sun angle and shadow reduces contrast which is essential for point identification in the SIFT algorithm. As Westoby et al. (2012, 302) stated "it is impossible to offer explicit guidance on the minimum number of photographs necessary for successful scene reconstruction". Here, the authors were referring not only to sun angle but also the variation in material and structure of features contained within any scene, all having an effect on the quality of reconstruction. Given a limited flight time, efforts were made to increase the redundancy of overlapping imagery to potentially increase the quality of the modelled

area. Completely nadir image dataset were practically impossible due to the movement of the aircraft and the holding of the camera by hand which was evident in a post review of the data. The post-capture inspection of the images revealed that, despite the majority of the imagery being taken with an 18mm focal length, the autofocus ranged the focal length between 18mm-55mm as part of the autofocus function.

The secondary data collected for LiDAR and the traditional photogrammetric technique (TPT) had enough coverage for the entire study area leaving SfM model extents as the only spatially limiting factor for GCP collection. A single DSM could not be extracted for the whole study area due to its size and the flightpath creating insufficient data to connect each site. To create the individual extents, ortho-mosaics of the SfM domains were generated by sorting and processing the imagery using SfM. The SfM imagery was sorted into usable sites informed by areas/features included in each scene typically forming land use image datasets. To ensure quality of the datasets, images that contained errors or did not meet the necessary overlap requirements (Snavely et al. 2008) were omitted. Once sorted a roughly georeferenced ortho-mosaic was computed in Pix4D Mapper using the CoCT aerial imagery as a reference. The process of how these mosaics are generated is covered in Section 3.3.1. These mosaics served to demarcate the individual domains of each SfM-generated DSM. Figure 3.4 shows the extents of all the individual SfM-derived DSMs created in this study informing the GCP collection process. Image coordinates are shown to illustrate the camera locations used to create these DSMs.

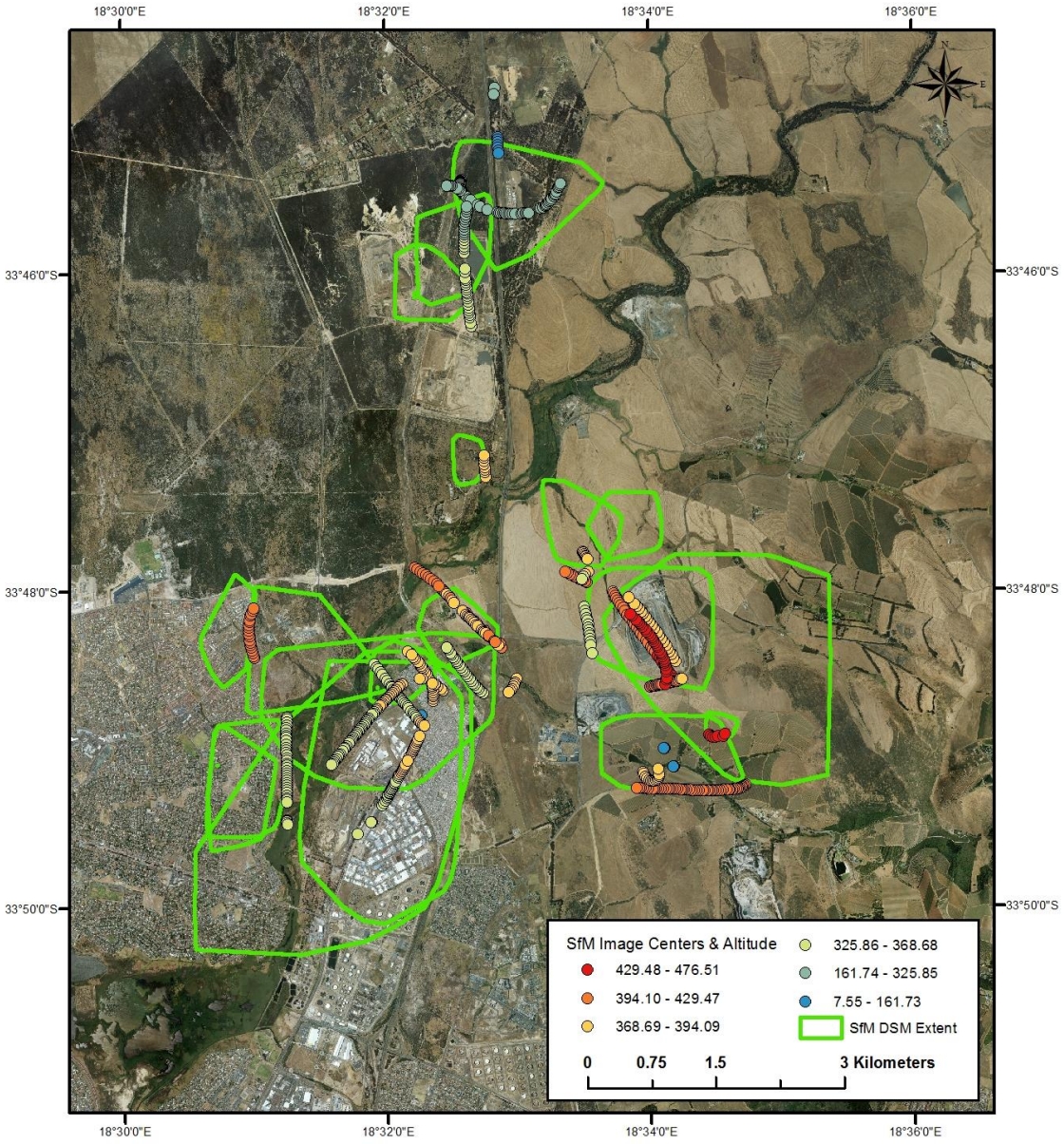


Figure 3.4 SfM DSM domains and estimated flying height at camera centres

The imagery collected covered a large area of which a large majority is private land, inaccessible or impractical for study. For the absolute accuracy assessment, the ground control points were taken manually as a reference with a differential GPS. In some cases access to areas was refused and therefore GCPs could not be captured. Typically, access was refused where safety was a concern (e.g. quarry), otherwise, where applicable safety, briefings and gear was provided. Reference ground control points were collected several months after the other datasets due to difficulties in negotiating access to informal settlements and other field areas. Some research areas were impractical for GCP collection (i.e. wetlands) as a result of fluctuating water levels and

vegetative growth making it unfeasible for identifying reliable ground control points and consistent height samples across all DSMs. These areas were removed from the study to ensure the validity of the results.

As no GPS points had been identified at this time, reference orthorectified aerial imagery was used for locational information. Stable features present on the orthophoto were collected as rough GCPs for the SfM data. Reliable ground control points were pre-identified based on roughly georeferenced SfM ortho-mosaics which indicated the DSM domains. The SfM domains informed the containment and placement for GCP collection ensuring adequate distribution (Figure 3.5 C & D) for georeferencing and testing. A field manual containing these point locations together with contextual maps (orthophoto and SfM ortho-mosaics) were taken into each site to locate the pre-identified points on the ground and update new or moved GCPs. In the event that the GCPs were moved, unstable or were unable to be found in the field, a new one was chosen based on individual SfM domains in the field (Figure 3.4, Figure 3.5D) and updated into the field manual. Contextual maps of each site, present in the field manual, were documented with the relocation information of the newly identified GPS points. Ground level photographs were taken for validation and identification of the point captured (Figure 3.5A & B).

Figure 3.5 (A & B) Contextual placement image of ground control point and (C & B) Aerial context map of GCPs present in the field manual

A surplus of 10 points per site was taken, from which a purposive random selection of four to five GCPs were separated for the georeferencing of the SfM DSMs. The purposive random sampling method was chosen to ensure adequate distribution of GCPs across the SfM model. The remainder of the points were used as independent samples for assessing the absolute accuracy with a minimum of five GCPs per site. A total of 197 GCPs were collected across all accessible sites. A Trimble GeoXH 6000 differential GPS with Tornado antenna was used. The collected coordinate data was sampled with a minimum of 60 observations per point or with an increased sample size until a 10cm spatial accuracy could be determined.

3.3 DSM GENERATION

This section presents the processing schemas of LiDAR, stereophotogrammetric and SfM methods for DSM extraction in correspondence with the state of the art presented in Sections 2.1 to Section 2.3. All vertical information was converted to orthometic height, height above mean sea level (MSL), as a standard using the South African geoid model (Chandler & Merry 2010).

3.3.1 Structure from Motion

SfM processing was broken up into four major procedures for the construction of DSMs: GCP insertion, initial processing, point cloud identification and DSM creation (Figure 3.1). An added step was included for the creation of ortho-mosaics used to inform GCP collection (Section 3.2.2). SfM processing is largely automated with little user input needed and was performed using Pix4D mapper software.

The subset of GCPs captured at each site was loaded into the software and images that effectively captured the corresponding object/feature points were manually tagged with this GPS information. This was done with the aid of the viewer provided for by the software. This process is not needed for 3D point cloud creation of the observed scene but without it, the point cloud is generated in relative coordinate system. This step informs the software of spatial location to georeference the point cloud. It is possible to have geo-tagged images where coordinates are recoded in the exchangeable image file (EXIF) format and upon reconstruction it can generate point clouds in a defined coordinate system without manual entry. This is a similar concept to the rational polynomial coefficient (RPC) files (with less information regarding external orientation) in aerial image processing with traditional photogrammetric techniques (TPTs).

The initial processing of SfM comprised of feature matching and 3D reconstruction in a single automated step in Pix4D mapper.

The images went through a key point identification process using SIFT (Lowe 2004). In this system features are identified in the form of candidate points and matched to candidate points identified in other images. If they are successfully matched by a minimum of three images, that point is no longer a candidate but identified as a key point. If an image does not contain the minimum of two key points, validated by two other images, that image is discarded automatically. The key points and their descriptors are passed to the sparse point cloud collection process estimating the camera orientation, performing bundle adjustment and reconstructing the scene (Figure 3.6A & C). The sparse point cloud is enhanced with more detail by decomposing subsets of the imagery and performing 3D reconstruction on each of these subsets, dramatically increasing the density of the point cloud in the final model (Figure 3.6B & D). The final DSM is constructed from a filtered point cloud, removing outliers and taking into account the surface features. While the exact details of Pix4D's filtering algorithm is proprietary information, the algorithm filters noise (erroneous points) by using the median altitude of neighbouring points. After filtering, erroneous 'bumps' are smoothed to create more linear surfaces, and allows for the user to adjust the sensitivity of this smoothing function. The smoothing function was set to 'Medium' which aims to preserve sharp features and flatten roughly planar areas. The software was set to recover a highest level of detail which did result in an increased processing time. SfM reconstructs a 3D scene (Figure 3.7) which has to be converted into a 2.5D DEM raster and, removing any information contained below, it is then possible to perform inter-raster calculations.

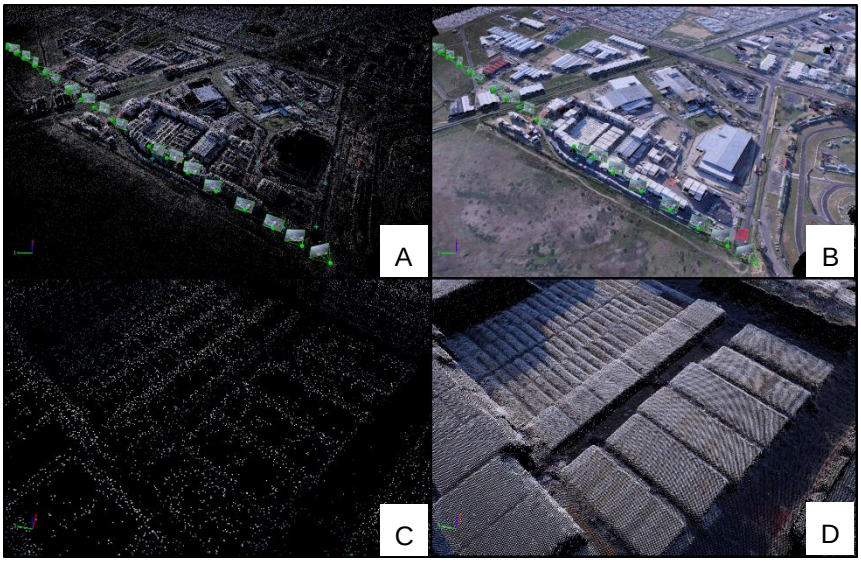


Figure 3.6 SfM processing of an industrial area (A) Sparse point cloud of scene (B) Dense point could of scene (C) Zoomed sparse point cloud within scene (D) Zoomed dense point cloud within scene

Figure 3.7 Flythrough of meshed scene of industrial area

Prior to the creation of the final surface models, roughly georeferenced ortho-mosaics were created to identify GPS locations for manual capture (Section 3.2.2). This was done to ensure the adequate distribution GPS GCP points. This process is part of the final outputs of Pix4D mapper along with DSM generation. Using the dense point cloud, imagery is stitched and ortho-rectified to create an ortho-mosaic.

3.3.2 Traditional Photogrammetry

The procedure used for computing the DSM using stereophotogrammetry followed three main processes: aerial triangulation, mass point production and interpolation. The images were processed into DSMs using Geomatica's Orthoengine 2014 (Table 3.2).

Aerial triangulation requires information about the interior orientation of the camera provided by the vendor (Table 3.1). Raw imagery and reference aerial data was collected from the CoCT's Geospatial Services (Table 3.2). The RPC files were provided by the CoCT which contain third degree polynomial coefficients that relate to the sensor's internal geometry and external geometry at the time of image capture. Additionally a manual GCP collection methodology was employed to increase accuracy of DSM over smaller areas involving only 2-3 images.

Reference orthophotos and a DEM were used to collect GCP information on stable features and updated to the imagery. Points were distributed evenly where possible with nine points per image. GCPs were also used as tie points, where possible, in the overlapping areas of each stereo pair, this ensured that at least two GCPs were collected as tie points. After manual collection, automatic GCP's and tie points were generated with each one validated against the reference. The points were inspected to ensure their precision, making manual modifications or deletions to the points where necessary. The software automatically computes triangulation error. This error was used to inform whether a point needed modification. The total error of the points had to maintain a sub-pixel root mean square error (RMSE) before processing could proceed. Subsequently, bundle adjustment was performed and epipolar images were generated to refine the exterior orientation and complete aerial triangulation of the image dataset.

To generate the individual DSMs, mass points were collected and interpolated to generate the surface. In Orthoengine, the DSM generation parameters were set to recover a high level of detail in a hilly environment which is contextual to the study site. A Natural Neighbour interpolator was used as a constant interpolator for TPT and LiDAR DSMs (Section 3.3.3). The DSMs were inspected for any observable reconstruction errors (e.g. artefacts and extreme discontinuities). If reconstruction error was found, the control and tie point were modified/recreated and the model rerun iteratively until a satisfactory result was acquired.

3.3.3 LiDAR

The processing of LiDAR is a well-documented technology and is expected to generate the most accurate DSMs owing to its intrinsic nature of range finding (Siebert & Teizer 2014). The LiDAR processing schema involves two main steps to create DSMs: point cloud filtering and interpolation. The LiDAR data was captured by the City of Cape Town (CoCT), Department of Geospatial service (Table 3.1).

The aerial LiDAR came in the form of an already pre-processed point cloud. The multiple returns had already been categorised only requiring the filtering out of the first returns from the dataset. The first returns were separated out to extract only the surface information to generate the DSM.

The detail can be increased with the use of a Natural Neighbour (NN) interpolator which will also produce the DSM raster. This method was employed to generate the DSMs used for study owing to its reconstruction abilities (Ekhatari et al. 2008; Bater & Coops 2009; Zhou 2013) which produces the best output with the least amount of human interaction (Bater & Coops 2009). NN is a deterministic approach utilising Thiessen polygons to weight the proportion of overlapping polygons for point interpolation (ESRI 2012).

3.3.4 DSM inspection

DSMs generated from three different techniques in three different software packages require data inspection as a qualitative measure to ensure they are compatible and to identify any errors in DSM generation. A manual inspection process was performed by overlaying all DSMs in ArcGIS ArcScene with a consistent colour ramp and a set elevation range to detect irregularities to address reconstruction error before analysis took place.

Figure 3.8 illustrates clipped DSMs created by TPT, Lidar and SfM across a residential, agricultural, informal settlement and industrial land use areas. The land use classes are confined to areas of interest that are used in accuracy evaluations. This is to limit any reconstruction produced towards the edge of DSMs which are prominent in image-based techniques.

Orthometric height was used as a standard across all models to ensure no offset due to vertical reference differences (Chandler & Merry 2010). Artefacts were discovered particularly in the TPT models with large vertical discontinuities in vine rows located in the agricultural areas which prompted refining and the recomputation of the surface.

Surface models that could not be adequately computed by all three methods or areas with large temporal variability identified by the DEMs were discarded.

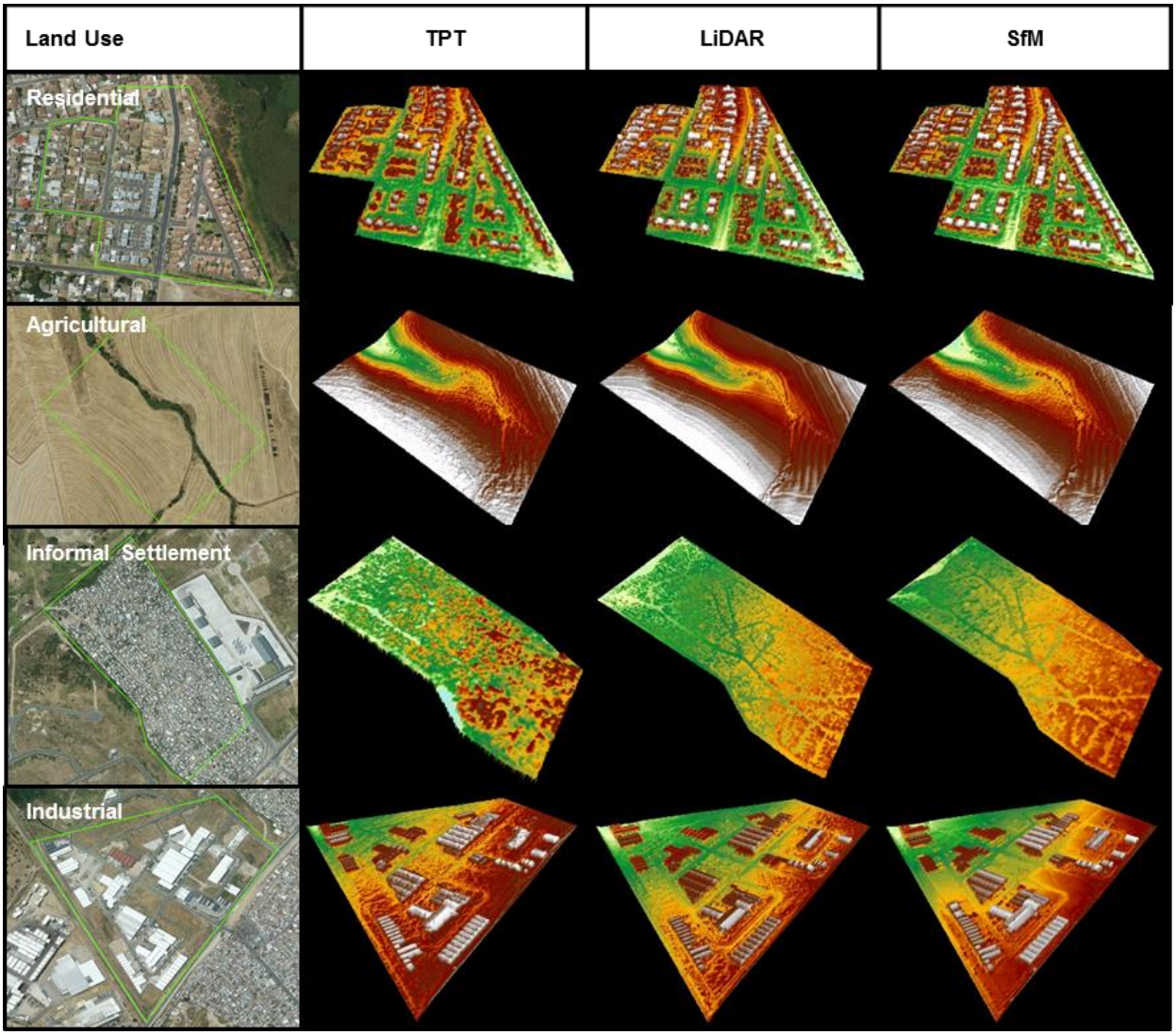


Figure 3.8 Land use area samples with resultant DSMs generated from TPT, LiDAR and SfM

Figure 3.9 illustrates the unclipped DSMs and their extents across all three surface generation techniques. The TPT extents may appear regular in this figure but these are the extents of the original uncorrected aerial imagery. These footprints were used as defining extents during orthorectification and as a result the outer edges DMS were written as null values where applicable. Due to the irregular shape of imagery based reconstruction, clipped extents (e.g. Figure 3.8) were used for processing to ensure adequate quality and coverage by all techniques.

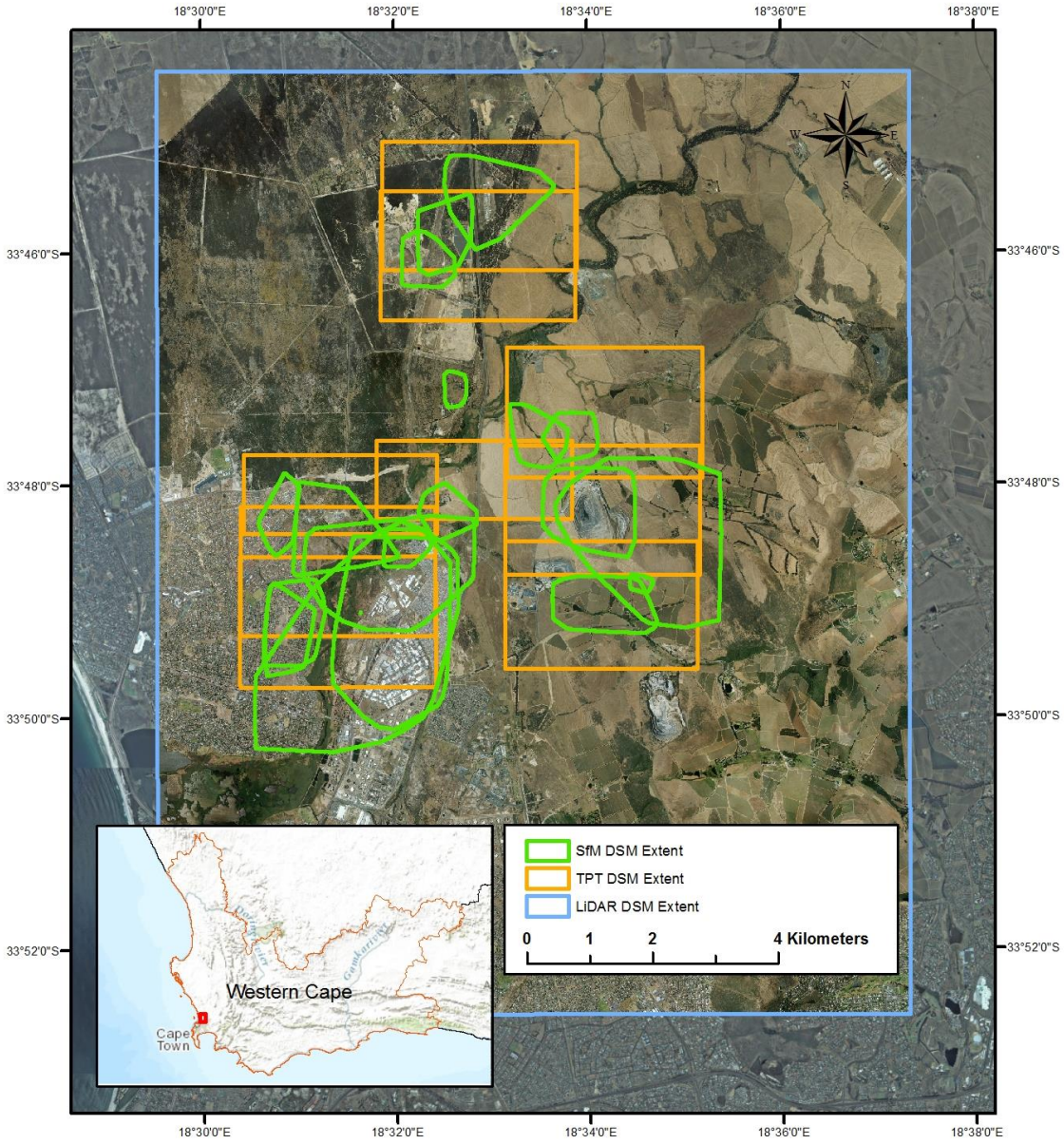


Figure 3.9 SfM, TPT and LiDAR surface model extents

Once all DSMs passed quality control, the DSMs were resampled to one meter spatial resolution. This was determined by the lowest resolution recorded of all DSMs created. This was done in order to calculate statistics on equally sized pixels for comparison and to limit bias from spatial resolution in the accuracy results. The final DSMs for all tests sites could then be evaluated through both qualitative and quantitative measures, as discussed in Section 3.5.

3.4 LAND COVER CLASSIFICATION

Effective land cover statistics for each land cover class is necessary to understand the nature of DSM reconstruction in different surface types and to assess reconstruction accuracy. This is in contrast to land use areas to which each site has been grouped. The classification was performed on the SPOT 5 multispectral satellite data with a national DEM (SUDEM) (Van Niekerk 2012) and the NGI 50 cm orthorectified aerial images as spatial references.

The land cover classification process was broken up into two major components: pre-processing and classification. The satellite imagery had to be pre-processed to remove any spatial and atmospheric error associated with the imagery. The classification process classifies both pixel and segmented pixel blocks in a hybridised approach.

3.4.1 Pre-processing

The software package used for pre-processing was PCI Geomatica's Orthoengine and Focus 2014 (Table 3.2) for orthorectification and atmospheric correction respectively. The panchromatic and multispectral images were orthorectified using NGI aerial imagery and a high resolution 5m national DEM (SUDEM) (Van Niekerk 2012). With both images orthorectified and coregistered atmospheric corrections were applied using ATCOR2. To limit any possible overcorrections due to relief, ATCOR2 was chosen over ATCOR3. To increase spatial resolution, the multispectral image was pan-sharpened using an automatic image fusion algorithm developed by Dr. Zhang (2002) and implemented in Geomatica Focus 2014. This is a subtractive method to recover structural information about objects in SPOT 5 imagery from the panchromatic image while maintaining spectral information across all four spectral bands.

3.4.2 Classification

Classification was performed in a hybrid method utilising two classifiers to extract different class information. A Support Vector Machine (SVM) was used to extract classes with typically high geometric and textural information while an Iso-Cluster classifier was used to separate spatially smaller and spectrally dissimilar classes. Although unconventional, this method integrates sound scientific classification principles (Campbell 2006; Myburgh & Van Niekerk 2013).

Object Based Image Analysis (OBIA) was used to segment the now pre-processed SPOT 5 image using Trimble's Econgintion software. The segmentation was initialised with parameters defined

by Myburgh & Van Niekerk (2013). These parameters were iteratively refined using visual inspection until objects were slightly over segmented. The adjustments to the segmentation parameters were made to identify more discrete boundaries with a finer scale of 8. Shape and compactness parameters, however, remained at 0.2 and 0.3 respectively. A support Vector Machine (SVM) algorithm was used to classify the image using 30 representative samples for each of the initial 10-16 pre-identified land cover classes. The classes were chosen based on the CSIR Land Cover Field Guide (2010). Table 3.4 shows the input parameters used in the SVM classifier. The inclusion derived datasets, such as NDVI and SAVI, were included to aid the classifier in class separation.

Table 3.4 Object features for SVM classifier

Type	**Measure**	**Feature**
Spectral Features	Mean Standard Deviation Ratio	Green, red, NIR, SWIR, brightness Green, red, NIR, SWIR Green, red, NIR, SWIR
Vegetation Indices		NDVI, SAVI
Texture Features	GLCM	Homogeneity, contrast, dissimilarity, entropy, ang. 2nd movement, correlation, mean, std. deviation
Geometry		Asymmetry, compactness, density, elliptic fit, radius of largest enclosing ellipse, rectangular fit, roundness, shape index

An accuracy assessment was performed using 400 (16 classes x 20 point per class + 80 point buffer) randomly generated points over the scene with the aim of a minimum of 10-15 stratified samples to compensate for the spatially smaller classes. This did not achieve adequate sampling of each class. Purposive sampling was undertaken of affected classes raising the total number of samples to 471. The values of the classification were extracted and pivoted against the samples to generate a confusion matrix and kappa statistic (Congalton & Green 2008). The overall accuracy was 62.85% with a kappa statistic of 0.597. This is insufficient to draw any meaningful conclusions from the classification despite a select number of classes having near perfect user's accuracy due to the high textural and geometric information (Congalton & Green 2008). These classes were extracted from this classification to be super imposed onto a pixel based classification.

An Iso-Cluster classification was used to separate spectrally dissimilar and spatially smaller classes (Campbell 2006). This was employed to limit spectral mixing and spectral homogeneity in segmented objects. From the unsupervised classification, classified pixels were grouped into meaningful land cover classes. These, in combination with the classes extracted from the SVM classification, were amalgamated to produce the final classification. Table 3.5 highlights the final classification schema and describes each of the 11 land cover classes used in this study.

Table 3.5 Classification schema

Land Cover Class	Description	Example
Artificial Bare Ground	Disturbed earth as a result of mining activity	Quarry
Bare Field	Sparsely vegetated area such as grasses and weeds due to recent harvests or lack of nutrition	Fallow Fields
Bare Ground	Bare soil as a results of environmental or human factors	Construction clearings
Hard Surfaces	Tar, concrete or gravel covered features	Roads
High Density Residential	High clustering of individual buildings per 100m x100m area	Informal Settlements
Medium Density Residential	Medium clustering of individual buildings per 100m x100m area	Urban Buildings
Industrial	Low clustering of large individual buildings per 100m x100m area (Industrial areas)	Industrial Buildings
Low Vegetation	Nutritionally high natural vegetation	Riparian vegetation
Trees & Shrubs	Natural broad canopy vegetation	Trees
Vineyards	Agriculturally prominent Crop in area	Vineyard Field
Water	Both natural and man-made water bodies	Dams

This method ensured higher classification accuracy by incorporating the strengths of OBIA and SVM, utilising textural and geometric features while not constraining subtle variance to larger generalized objects created by segmentation. Unsupervised classification constrains the spectral information to the size of a single pixel, limiting spectral mixing. This hybrid method seeks to take advantage of the strength of both classifiers.

3.4.3 Accuracy assessment of land cover classification

The adoption of a hybrid method yielded favourable confidence in the land cover classification (Table 3.6) with a high level of overall accuracy (88.95%) and a strong level of agreement with a

KHAT statistic of 0.86 (Congalton & Green 2008). The resulting classification dataset (Figure 3.10) could then be treated with a high level of confidence.

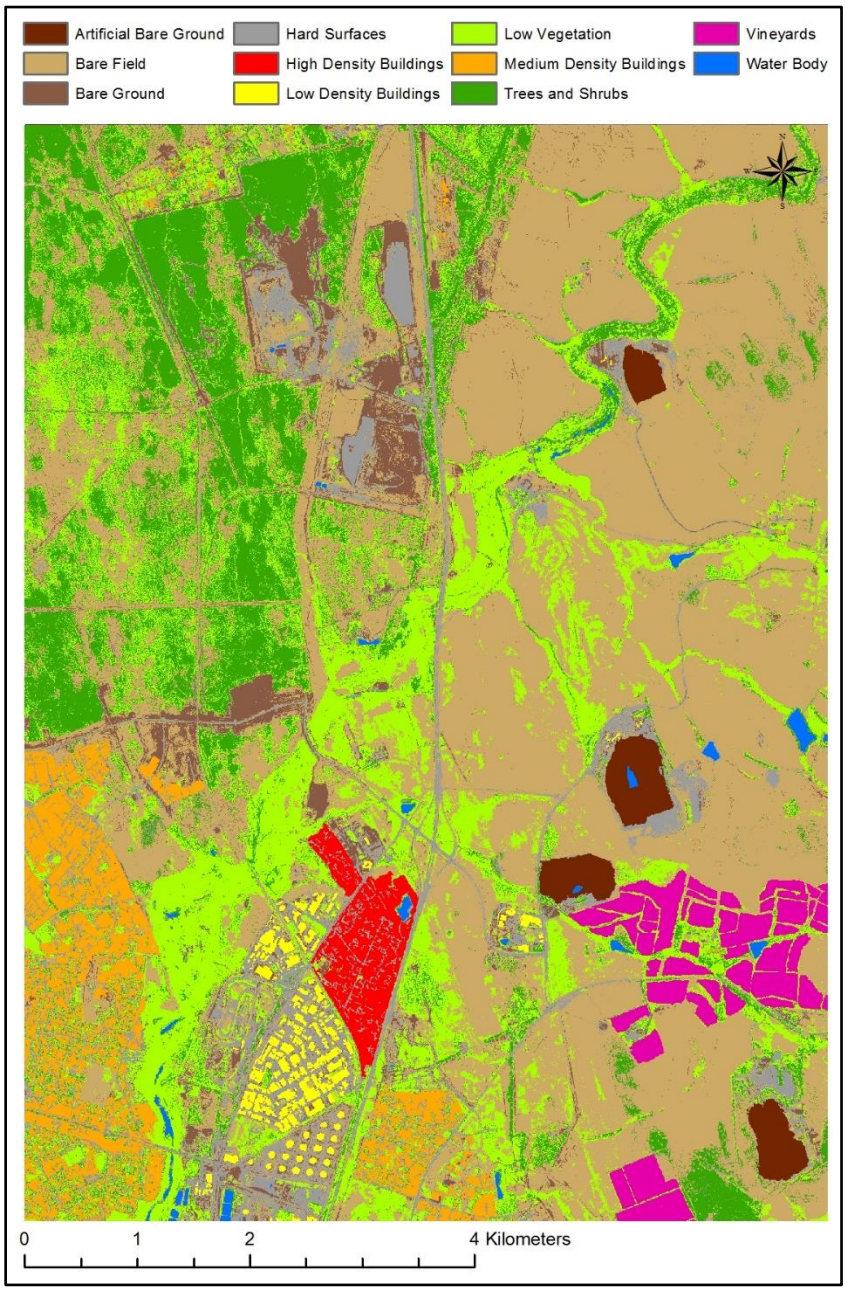


Figure 3.10 Classification of study area

The average producer's accuracy (93.34%) indicates a strong probability of correct classification and average user's accuracy (90.05%) with a high likelihood this classification is an accurate reflection of the ground truth.

Table 3.6 Confusion matrix evaluating classification accuracy

	A	B	C	D	E	F	G	H	I	J	K	Total	Users	Omission
A	17											17	100.00	0.00
B		132		2			5		1			140	94.29	5.71
C		2	22	1								25	88.00	12.00
D		2		28								30	93.33	6.67
E		1	1	2	21							25	84.00	16.00
F						29	1					30	96.67	3.33
G		10					70		5			85	82.35	17.65
H				2			1	24				27	88.89	11.11
I		3					11		38			52	73.08	26.92
J										20		20	100.00	0.00
K		1		1							18	20	90.00	10.00
Total	17	151	23	36	21	29	88	24	44	20	18	471		
Producers	100.00	87.42	95.65	77.78	100.00	100.00	79.55	100.00	86.36	100.00	100.00			
Commission	0.00	12.58	4.35	22.22	0.00	0.00	20.45	0.00	13.64	0.00	0.00			

Accuracy	**88.96**
KHAT	**0.86**
Producers	**93.34**
Users	**90.05**
Omission	9.94
Commission	**6.65**

A	Artificial Bare	**G**	Low Vegetation
B	Bare Field	**H**	Medium Density Residential
C	Bare Ground	**I**	Trees & Shrubs
D	Hard Surfaces	**J**	Vineyards
E	High Density Residential	**K**	Water
F	Industrial		

The highest levels of commission error is the 'Hard Surfaces' class (22.22%) which contains composites (tar, gravel, concrete etc.). It is also present in other classes such as buildings and 'Bare Ground' is confirmed by the confusion matrix. The main source of omission is the vegetative 'Low

Vegetation' class (17.65%) which has strong inclusion in 'Bare Field' and 'Trees & Shrubs' that typically boarder these area and can spectrally exhibit similar spectral responses.

3.5 ACCURACY ASSESSMENT

The objectives set out by this study were to assess the vertical accuracy of DSMs developed by LiDAR, SfM and stereophotogrammetry to a common area to an independent reference and perform a comparative relative error assessment between the derived data. To achieve these objectives, two accuracy assessments were performed: absolute and relative.

Figure 3.11 illustrates the workflow of accuracy assessment process. The absolute accuracy assessment comprised a single assessment that measured the vertical difference of every DSM to the independent GPS points collected in the primary data capture process (Section 3.2.2). These accuracies were then aggregated into a single accuracy for each land use class. The relative accuracy assessment was broken up into five sub-assessments dealing with specific components covered in this section. These assessments look at the inter-relational performance of each DSM technique to each other in surface reconstruction.

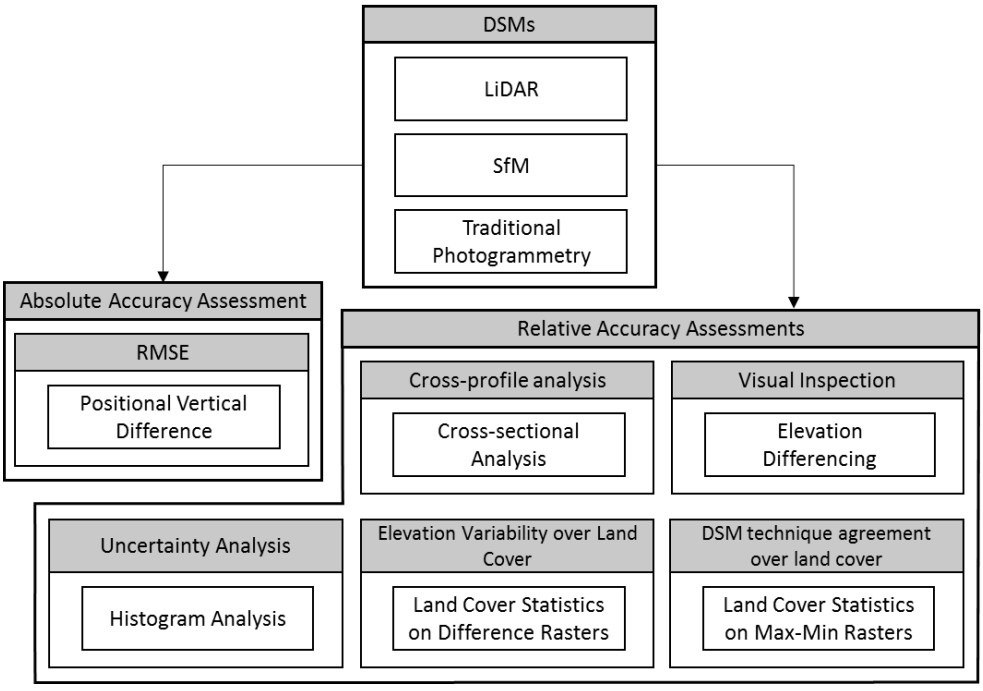


Figure 3.11 Absolute and relative accuracy assessments evaluating the datasets

3.5.1 Absolute Accuracy Assessment

Absolute error was calculated by measuring the vertical offset of each surface model from the reference GCPs. Due to the nature of elevation models, assessment in the X and Y coordinates were not possible as no visible information was present. This is a result of points having been typically captured based on visible and identifiable markers (i.e. lines on roads) which may have no vertical difference therefore making them unidentifiable in the LiDAR DSM. The GCPs used for the accuracy assessment were collected in the same process as the GCPs used for the georeferencing of SfM DSMs. To ensure consistency and validity, ground control points used to reference SfM models were not included in the assessment.

The accuracy assessment is the level of agreement between the DSMs and the field collected GCPs and is aggregated based on land use. The land use classes are divided into four categories based on predominant land use of the clipped regions of the area: industrial, residential, informal settlement and agricultural areas. The vertical accuracy was measured by calculating the Root Mean Squared Error (RMSE) and averaging the results per land use area. RMSE is a scale-dependent measure of the difference (or residuals) between the actual and predicted values in a dataset evaluated by squared error or quadratic loss. This function produces a value representative of the magnitude of error.

3.5.2 Relative Accuracy Assessment

Relative error observation was approached using multiple methods to describe the nature of the resulting data (Figure 3.11). Although the absolute error assessment indicates vertical accuracy at specific points, it does not take into account variability in land cover since stable reference points for a DSM cannot be established on all objects (i.e. top of vegetation, roofs, water etc.). This necessitated the use of relative methods. Five methods of relative accuracy assessments were performed each assessing the quality of DSM creation: visual inspection, cross-profile analysis, elevation variability over land cover, DSM technique agreement over land cover and uncertainty analysis (Figure 3.11). Specific attention was paid to SfM as the aim of this study was to review and contextualise this method with regard to existing and well established techniques.

In some cases, analysis was performed with difference rasters where DSMs were subtracted either by extrema or pure elevation values (Agugiaro et al. 2012). There are two main types of difference rasters used for testing: simple subtraction and local extrema subtraction.

Simple subtraction used DSMs from the most accurate DSM construction method, LiDAR (as identified by the absolute accuracy assessment), as base raster to which SfM and TPT DSMs were individually subtracted. This LiDAR base layer remained static to ensure consistency for the duration of the study. This served to describe the deviation from the superior technique to identify where deviation existed.

Local extrema subtraction was created from two rasters: local minima raster and local maxima raster. The local minima raster was created by taking the lowest elevation value recorded across all three DMSs (LiDAR, TPT and SfM) per unit area. These minimum elevation values produced a surface raster that recorded all the lowest elevations observed. This process was mimicked by the maxima raster with the exception that it recorded the maximum values observed. The minima raster was subtracted from the maxima raster to produce a raster that illustrated areas with the highest level of agreement and non-agreement across all three DSM techniques used in the study.

3.5.2.1 Visual Inspection

Visual inspection refers to error that can be visibly identified both during DSM production and in the final surface models. Further error was identified by using the simple subtraction raster to detect artefacts and discontinuities that may have been generated during DSM creation. These irregularities were catalogued and analysed in the context of offset, reconstruction error and variance. This analysis provided a comparative description into the nature of some of these errors as considerations when implementing these techniques.

3.5.2.2 Cross-profile Analysis

Cross-sectional profiles provide structural and displacement information across all three datasets over specific two dimensional transects (Dietrich 2014; Smith & Vericat 2015). Sites were selected in purposefully random manner based on discontinuities observed in the differenced rasters. Linear segments were traced over these areas with an elevation sample taken every 10cm. These tests will form errors in reconstruction typically associated with over-smoothing, variation in contrast, estimation error and displacement. The profiles will provide information regarding surface recovery which may not be evident from a vertical perspective.

3.5.2.3 Elevation variability over land cover

Elevation variability uses statistical methods to calculate the minimum value, maximum value, range, mean and standard deviation across over each land cover. This assessment was computed on the simple difference rasters individually. The statistics were computed for each land cover class (Section 3.4) to describe the vertical accuracy associated with each land cover class. Study sites were clipped to ensure the measurements did not include error resulting from deformations on the perimeter of the DSMs. It is important to note that not all land covers were represented in each study site. The statistics were individually calculated and aggregated for local and global analysis respectively.

3.5.2.4 DSM technique agreement over land cover

Similarly to elevation variability over land cover, DSM technique agreement over land cover computes the same statistics with the substitution of the local extrema raster. Agreement was observed by extracting the difference of absolute maximum and minimum elevation value per unit area across all three DSM generated by each technique.

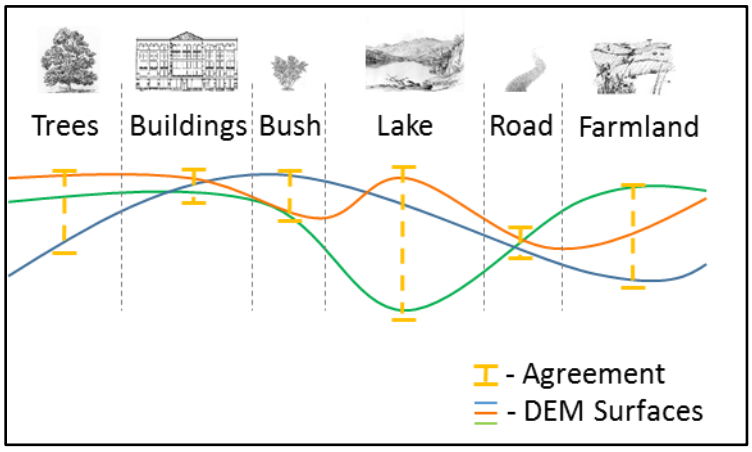


Figure 3.12 Example of agreement measured by the range defined by the extreme minimum and maximum (extrema) values observed by all three DSMs created over an area

Figure 3.12 visually depicts the agreement of each DSM development technique and how well they agree over land cover areas. This measurement technique illustrated the areas subject to the largest level disagreement between all techniques across land cover classes. Land cover classes with large variation indicate disagreement between techniques about the surface and the converse is true with low variation. This identified which land cover areas needed more consideration in surface reconstruction methods.

3.5.2.5 Uncertainty Analysis

Vertical uncertainty or roughness (Clapuyt et al. 2015; Smith & Vericat 2015) is the graphical and statistical description illustrating the degree of vertical error distribution of the data within the simple difference rasters. The graphical descriptor is a histogram that illustrates the inter-DSM correlation and vertical error distribution in specific land use areas. This was done for each land use class for LiDAR – SfM and LiDAR – TPT difference rasters. Statistical descriptors coupled with the histograms described how closely the data matched that of LiDAR and if that representation was consistent. A normal distribution cannot be expected due to aggregation of the data.

CHAPTER 4: RESULTS & DISCUSSION

This chapter presents and interprets the results in terms of the relative and absolute accuracy assessments. The objectives of this study was to identify limitations of, and evaluate suitability for, and vertical accuracy of, each method in DSM extraction. Feasibility considerations are also added to identify external factors affecting each technique in conjunction with the accuracy assessments. A focus is placed on SfM as the aim seeks to compare this DSM technique with well-established techniques in varying environments.

4.1 ABSOLUTE ACCURACY

Out of a total of 24 candidate DSM sites, only 13 were found viable due to site access, safety, surface change (mining activity, flooding, vegetative growth) and spatial overlap between DSMs (Section 3.2.2). The remaining sites were grouped into one of four land use types: Informal Settlements; Residential; Industrial and Agricultural Land. The absolute accuracy assessment depicted RMSE values of all vertical offsets from GPS points across all 13 sites for each of the three DSM types. Figure 4.1 shows the average RMSE calculated across all 13 sites.

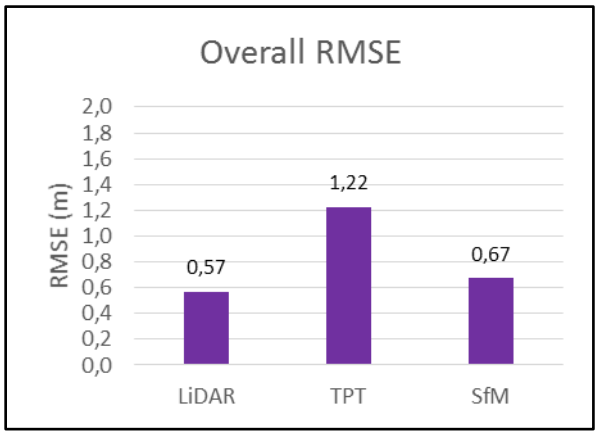


Figure 4.1 Average RMSE for all surface models (n=13) measured against differential GPS reference points

According to literature (Table 2.1) we expected that aerial data collection methods with a GPS reference should produce a SfM vertical error range of 0.07-0.37 m (Javernick et al. 2014; Hugenholtz et al. 2013; Fonstad et al. 2012) and traditional image-based vertical error between 0.31-0.58 m (Ahmad et al 2012). Including Total Station Surveys as more accurate equivalents to GPS measurements, the vertical error range can be extended to 0.07-0.59 m for SfM based assessments (Tonkin et al. 2014; Harwin & Lucleer 2012). The lowest overall average RMSE

confirmed that LiDAR was vertically the most accurate (Javernick et al. 2014; Hugenholtz et al. 2013) achieving 0.57 m total error, followed by SfM at 0.67 m and TPT at 1.23 m. These results were as anticipated as LiDAR is largely a range-finding technique and not an estimation technique based on epipolar geometry (Siebert & Teizer 2014).

LiDAR performed particularly well in urban environments (Figure 4.2A & B) achieving the lowest RMSE in both informal (0.29 m) and formal residential areas (0.17 m), followed by SfM (0.63 m, 0.48 m) and TPT (1.35 m, 1.24 m). The proximity of objects in urban areas and the clustering of informal dwellings may result in overgeneralization in TPT and the failure of SfM in the meshing process, resulting in higher RMSE values. LiDAR, however, is not as limited by the proximity or clustering in urban areas as ranged data is measured at set intervals regardless of what features are being sensed. The average error measurement for SfM across all six urban sites (residential and informal settlement area) was 0.53 m with a vertical error range of 0.16-0.76 m between DSMs, whereas the error for LiDAR was 0.21 m with an error range of 0.13-0.46 m.

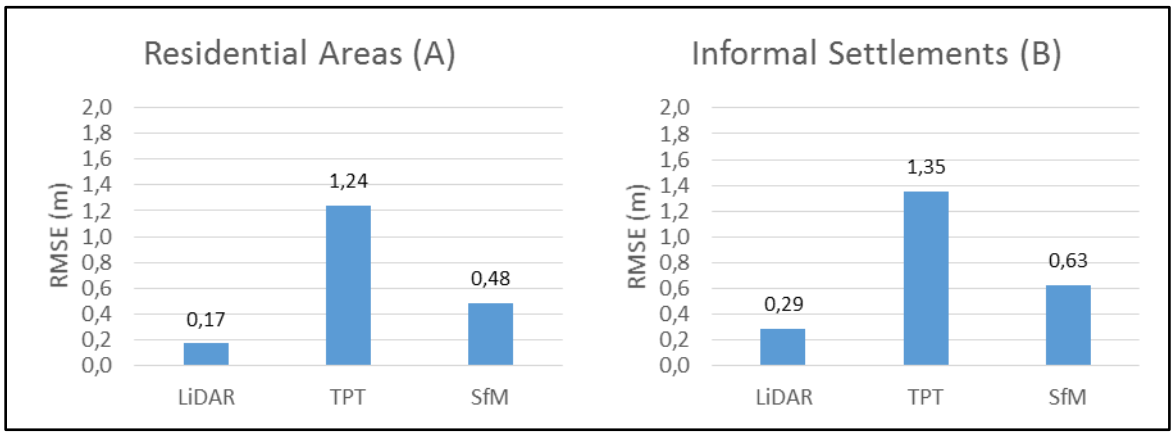


Figure 4.2 Average RMSE for residential areas (A) and informal settlements (B)

Further urban data exploration revealed the disparity observed between LiDAR and SfM was 0.31m. This SfM-LiDAR margin is in line with accuracies found with aerial based collection of SfM imagery with a LiDAR based reference (Fonstad et al. 2012; Stal et al; 2012 Johnson et al. 2014). Using terrestrial methods Dietrich (2014) found results around 0.18-0.20 m range illustrating a significant deviation from the airborne platform approach used here.

SfM better represented areas in environments that have more linear, larger and easily definable objects such as land fill foundations and industrial areas. The absolute accuracy can be seen in Figure 4.3 illustrating the average RMSE values recorded over industrial areas. SfM error in these DSMs were in the vertical error range of 0.13-0.28 m across all four industrial areas, achieving an

overall RMSE of 0.28 m. It was also observed that traditional photogrammetry performed particularly poorly in industrial areas with a vertical error range of 1.1- 2.2 m and a RMSE of 1.83 m. LiDAR performed poorly compared to SfM in these environments by a RMSE deviation of 0.37 m with a final RMSE of 0.58 m and vertical error range of 0.50 – 0.76 m (Figure 4.3).

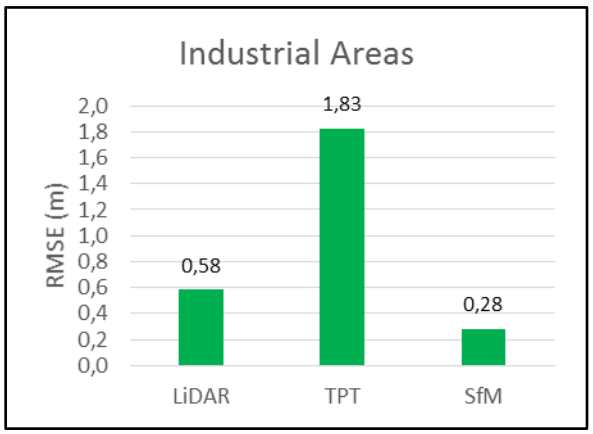


Figure 4.3 Average RMSE for industrial areas

An investigation into the reason for this finding revealed that temporal factors may have contributed to the weaker performance of LiDAR and TPT. A review of reference imagery confirmed that one study site containing landfill foundations was nearing completion at the time of stereo imagery and LiDAR collection whereas SfM imagery and GPS samples were captured post-completion with resulting structural inconsistencies within the scene. The highly textured nature of these environments, with easily definable large objects, may potentially be responsible for the superior SfM result. The results obtained in the industrial areas were consistent with the accuracies found by Javernick et al. (2014) (0.13-37 m) and Hugenholtz et al. (2013) (0.29 m) who used GPS-based references.

Traditional photogrammetry measured the lowest vertical error in agriculture areas (Figure 4.4). The generalized feature landscape and hilly terrain may have yielded better interpretations of the rendered surface of TPT in agricultural areas. The lowest RMSE in agricultural areas was recorded by TPT (0.52 m), followed by LiDAR (1.27 m) and SfM (1.35 m) as shown in Figure 4.4.

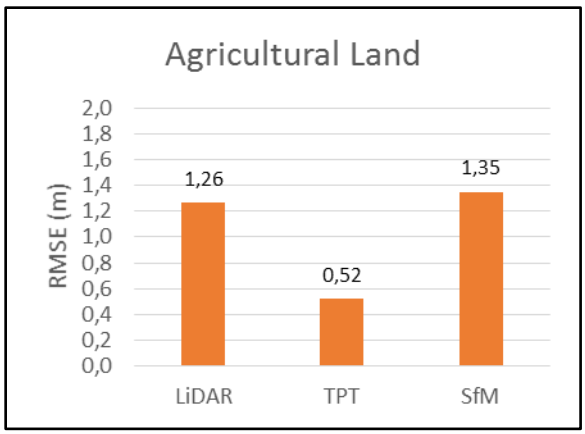


Figure 4.4 Average RMSE for agricultural areas

The large error for all methods is consistently high with a vertical error range of 0.76-1.7 m across all three agricultural areas. The SfM-LiDAR relationship demonstrated a 0.08 m error disparity margin confirming SfM's comparability with LiDAR (Fonstad et al. 2012; Westoby et al. 2013) despite demonstrating high vertical error. When considering the nature of this disparity, a temporal offset was ruled out as a possible cause of error as ALS data was collected in a similar time frame to that of the traditional photogrammetry data. Further investigation into the primary data revealed that GCPs were collected on hard ground and not on top of vegetation where interpolation techniques may result in over-estimations by points captured in close proximity to vegetation. This could potentially contribute to the error in LiDAR and SfM compared to TPT.

The 1.34 m error demonstrated by SfM is not atypical of the local vegetation height above ground level. It is not uncommon for SFM to have larger errors in vegetated areas (Javernick et al. 2014; Johnson et al. 2014; Tonkin et al. 2014). Westoby et al. (2013) observed a similar accuracy with 90% of the data within 1m of the reference surface using terrestrial methods. The influence of objects typical in agricultural areas is substantial with the impact of vegetative offset having marked effects. Had agricultural areas been excluded, the overall absolute error in SfM would decrease by 35.8% to a total RMSE of 0.43 m (down from 0.67 m) with a range of 0.11-0.94 m (down from 0.11-1.77 m) which is congruent with literature highlighting the impact of vegetation.

4.2 RELATIVE ACCURACY

The relative accuracy has been evaluated in several ways to describe the DSM reconstruction accuracy relative to each technique. Difference rasters form an integral part of the relative accuracy assessment (Agugiaro et al. 2012) with two main types of difference rasters having been created: simple subtraction raster and local extrema difference raster for analysis (Section 3.5.2).

There were five techniques used to assess relative accuracy: visual inspection, cross-profile analysis, elevation variability over land cover, DSM technique agreement over land cover and uncertainty analysis:

- Visual inspection identifies visible error during DSM production and in the final surface models.
- Cross-sectional profiles illustrate structural and displacement information over specific 2D transects across DSMs created by each technique.
- Elevation variability over land cover uses statistical methods to calculate the minimum value, maximum value, range, mean and standard deviation over each land cover.
- DSM technique agreement over land cover computes the same statistical method with the substitution of the local extrema raster to indicate levels of agreement between techniques.
- Vertical uncertainty is the graphical and statistical descriptions showing the degree of vertical distribution of error within the simple difference rasters.

No single accuracy assessment was sufficient to provide substantial and conclusive evidence of the success of the DSMs, hence multiple accuracy assessments were performed to identify shortfalls and potential weaknesses in DSM creation across the different techniques.

4.2.1 Visual Inspection

SfM's performance is based on the texture of the environment observed. Where little texture is present between two individual features, due to low contrast, errors in point creation can occur. Figure 4.5 illustrates points incorrectly identified on the edge of a plane hangar. The low contrast of the shadowed concrete floor had enough texture for a point to be identified but too little contrast to correctly estimate the point location. More points are mis-appropriated towards the centre of the

roof as less and less contrast is given by its surroundings as it approaches the centre of the concrete floor.

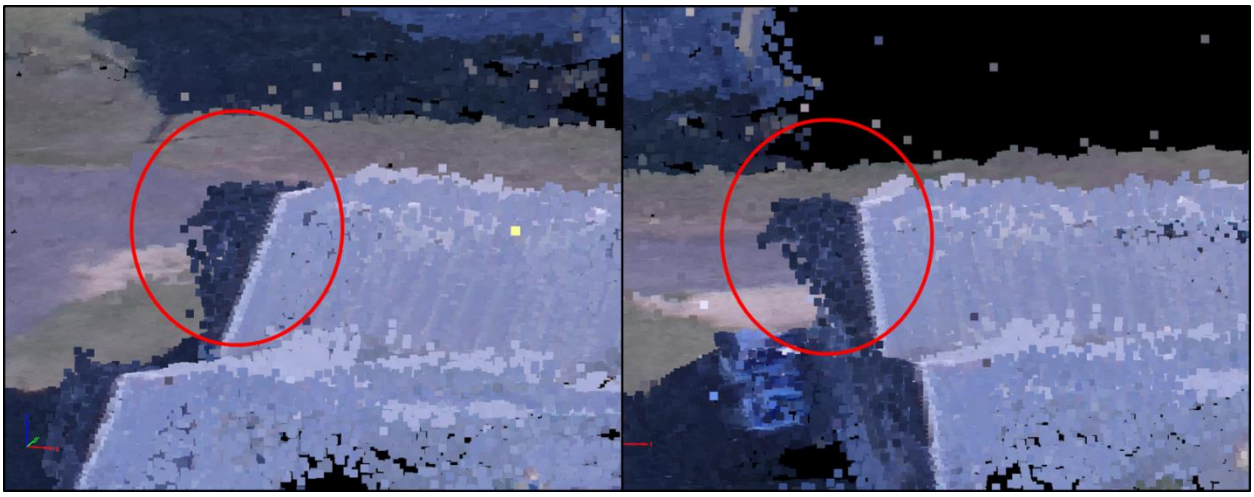


Figure 4.5 Incorrect point estimation of plane hangar roof due to low contrast from two view angles. Shadowed ground misinterpreted as part of the roof structure as highlighted by red container.

Typically this is a result of homogeneity of surface features, sun angle and shadowing effects causing the failure of point identification or the misappropriation of points around ground features (James & Robertson 2012; Bemis et al. 2014; Johnson et al. 2014). These artefacts are generally filtered out by removing erroneous points by using median elevation of neighbouring points or by the surface smoothing algorithm used by Pix4D Mapper. Jancosek & Pajdla (2011) use point occlusion to help with these weaker and more homogeneous surfaces. This application works best when points are created closely together on two homogenous surfaces (i.e. two dividing walls). The occlusion of points in one image and not in another from a different angle suggests that a surface exists between them, despite limited data being identified on the surface itself.

The use of extreme oblique imagery (extreme non-nadir image angles) can create large footprints on the ground, resulting in substantial detail loss toward the edge of the DSM. Figure 4.6 illustrates the simple difference raster created by LiDAR-SFM of an airfield defined by the SfM extent. The area of interest (A) is used for testing but significant error (red to orange shades) is observed in large sections of the SfM DSM extent (B). On the right side tessellation blocks are outlined (C) that show the low level of detail as a result of reconstruction with oblique imagery taken from extreme angles.

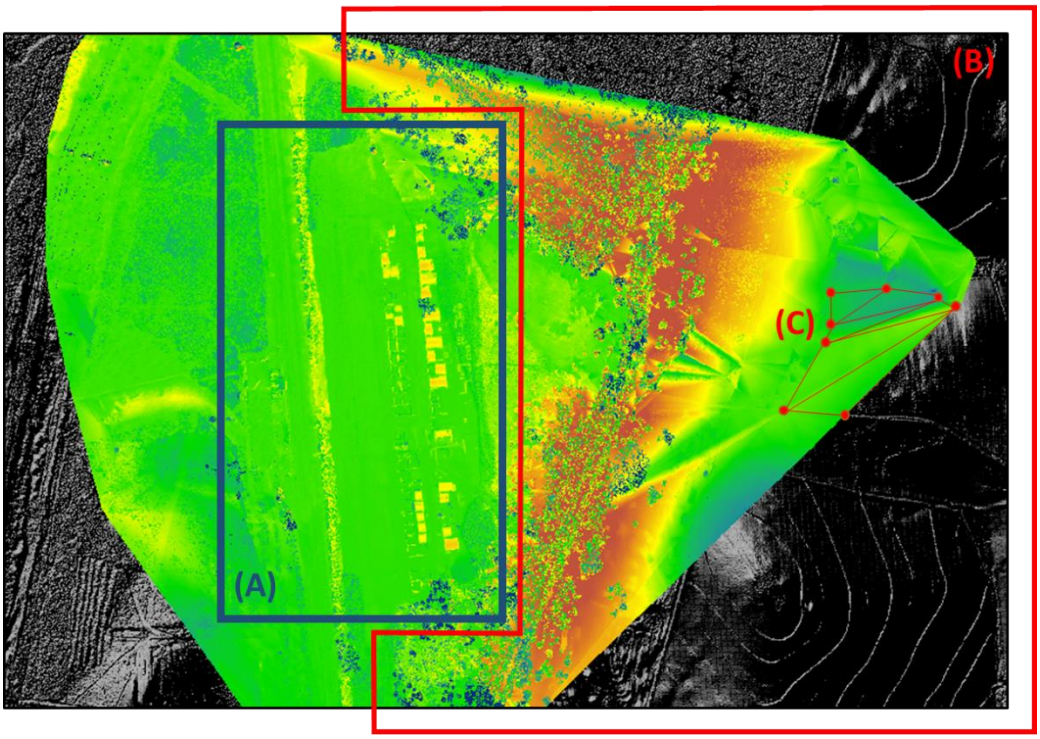


Figure 4.6 Full SfM DSM reconstruct of an airfield illustrating the (A) area of interest of airfield, (B) detail decay and overgeneralization of surface with (C) large linkages between points

This detail decay (Figure 4.6 B & C) resulted from imagery that contained a few stable key points located in the distance as satisfying the overlap criterion (Snavely et al. 2008). Large generalized areas were tessellated in what was effectively less detailed and less accurate representation of the surface (Johnson et al. 2014). These areas were inaccurate and could skew error metrics if they were included in an accuracy assessment where their inclusion may adversely affect the total accuracy with non-uniform distribution of error. It is then important to clip the extents to capture the relevant area needed for study to remove this discontinuity and gauge a better confidence of the area to be studied. As a result, it is recommended to collect imagery well beyond the study extent to ensure the integrity of the research area. With image capture this issue can be resolved by nadir and near-nadir imagery with a relatively small variation in capture angle variance. A convergent image configuration is recommended as oblique imagery provides meaningful information on the vertical profiles of objects in a scene which may be lost when sensing from a purely nadir position (Wackrow & Chandler 2008).

Figure 4.7 is an example of how ‘dead ground’ (Westoby et al. 2012) is created and the effects it can have when reconstructing a surface. From the oblique view of the house (right), no information is gathered about the left side of the building. This creates points only on the roof and at ground level leaving only an assumed gradient over the area with no data. Dead ground is the concept of non-sensed area due to occlusion and is particularly prevalent in terrestrial sensing from a single side of an object.

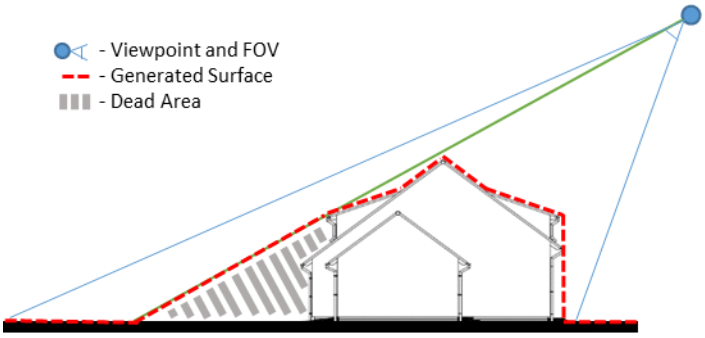


Figure 4.7 Illustration of dead ground produced by oblique imagery

If this oblique imagery is only captured from one angle, dead ground can negatively impact reconstruction (Figure 4.6 & Figure 4.7). Strikes and dips on a surface can produce this error if insufficient imagery in captured around the observed surface.

Figure 4.8 illustrates a practical example of dead ground in a residential area taken from a SfM-LiDAR difference raster. These areas are typically a conglomerate of housing, fences and vegetation, highlighted in pink. The red areas to the left of the houses in Figure 4.8A are surface generalisations created as a result of oblique imagery highlighted in blue in Figure 4.8B. These areas of no information can also be observed in the original SfM point cloud with the white background visible to the left of the homes highlighted in pink (Figure 4.8B).

Figure 4.8 Practical example of dead ground in a residential area created from oblique imagery, resulting in overgeneralization. (A) LiDAR – SfM simple difference raster illustrating vertical elevation deviation with homes highlighted in pink. (B) Dense SfM colour mapped point cloud over the same residential area with imagery location and angle highlighted in blue.

The pink areas in Figure 4.8 and red area in Figure 4.6B are overgeneralizations due to the lack of information of occluded surfaces. In both figures, tall objects (trees, housing, fences etc.) are the main contributors of occlusion in the SfM imagery. If a house restricts the view of a field behind it and no other information is present, the meshing algorithm can only assume a linear/curved connection between the roof and the nearest point collected on the ground (Figure 4.6B & Figure 4.8A). The over-generalisation of these areas also produces an over-estimation of the surface seen in the simple difference raster (Figure 4.8A).

In the South African case, this type of error may be largely unavoidable due to local aviation laws. Special permission has to be obtained to fly over residential areas at low altitude for commercial aircraft and UAV (remotely piloted aircraft systems (RPAS) in the South African context). Flights over these residential areas are not permitted unless signed permission is obtained from the director of the South African Civil Aviation Authority (South African Civil Aviation Authority 2015). In the case of aerial LiDAR and image data collected by the CoCT, the flying height is substantially higher, subverting the need for special permission, but resulting in the need for higher resolution sensors.

Despite TPT producing superior absolute vertical accuracy in agricultural areas (Section 4.1), it produced artefacts in vine rows necessitating iterative removal by refinement and reprocessing the DSM. This error is an over- and under-estimation of the geometry associated with the vine rows with vertical deformations producing sharp discontinuities of up to 22.6m. As a comparatively manual technique it is exacerbated by manual re-dispersion of GCPs to create an adequate output. These errors may not be detected using point-based accuracy assessment as a point has to be captured in these specific areas for them to be observed.

4.2.2 Cross-profile Analysis

Cross-profile analysis evaluates straight line samples of objects and features in selected areas to better describe the vertical quality and detail of reconstruction across all three techniques. Samples were taken at 10cm intervals along profile lines irrespective of the one meter DSM resolution to better capture vertical change from pixel to pixel.

Vegetated areas present particular difficulties with image-based reconstruction (Tonkin et al. 2014; Javernick et al. 2014; Johnson et al. 2014). The reconstruction of individual large trees and tree lines can be seen in Figure 4.9 & Figure 4.10 respectively. TPT presented an over generalized surface across both transects generally under-sampling of large vegetated features. In the instance of a single tree, the vertical profile was largely unrepresented. An offset is observed to the right of the individual canopy (Figure 4.9) which may result from the level of foliage observed at the time of capture. Gaps between individual branches created a highly variable vertical response from SfM where first return LiDAR captured the top of the entire canopy. This may be a result of the spatial resolution of the LiDAR points versus the spatial resolution with the level of contrast between branches in the SfM imagery.

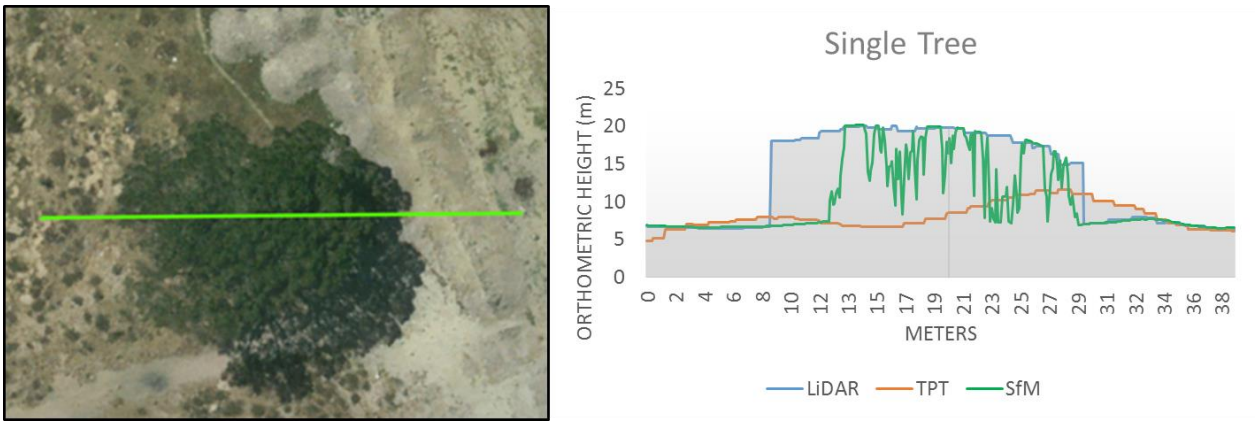


Figure 4.9 Transect of a single tree with image reference

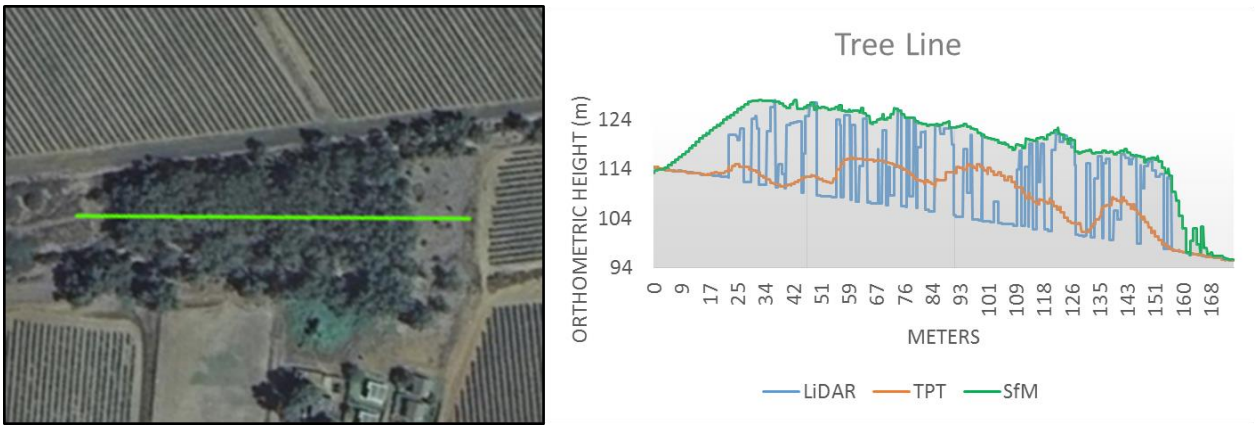


Figure 4.10 Transect of a tree line with image reference

The converse is true for the tree line where LiDAR was more sensitive to gaps in the tree line, and where SfM captured the whole of the canopy as one continuous surface (Figure 4.10). SfM also illustrated over generalized features at the ends of the tree line due to shadows which reduced contrast. SfM was able to identify vegetative objects but showed limitations in the separation of clustered features making delineation, and possible removal, daunting. This limitation must be considered in the conversion from digital surface model (DSM) to digital terrain model (DTM) if information about the terrain is to be extracted which can be a useful tool in GIS especially over larger areas (Van Niekerk 2012).

Figure 4.11 illustrates more regularly spaced vegetation in vine rows (with set distances between vines). Larger surface agreement between SfM and LiDAR was observed with regards to object structure. Despite the apparent offset, SfM did mimic delineations between vine rows and captured the contained vegetation.

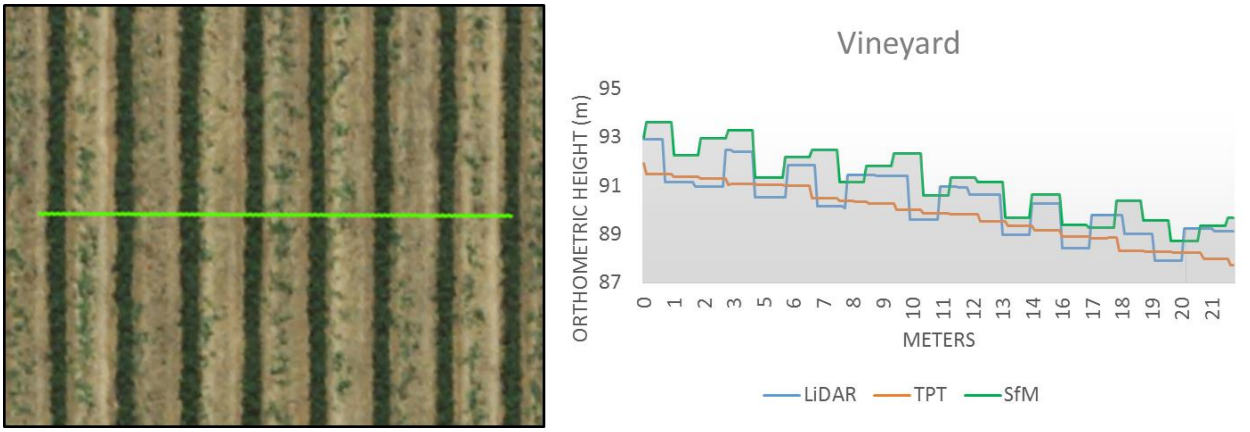


Figure 4.11 Transect of crop rows in a vineyard with image reference

TPT failed to recover vertical information of the vine rows in this cross-profile which is in contrast to the high vertical accuracy TPT demonstrated in the absolute accuracy assessment over agricultural land. The absolute accuracy (Section 4.1) measured the lowest RMSE value recorded for TPT (0.52 m), followed by LiDAR (1.27 m) and SfM (1.35 m). The overgeneralization of agricultural land is believed to be the cause of the lower error as sample GPS points were captured on terrain-level features. As vegetation is largely unrecorded by TPT, the terrain is better observed than the surface. As a result, TPT had a higher likelihood of predicting the position of the control point on the ground. Furthermore, the interpolation of LiDAR and the filtering of points can create a smoothing effect when transitioning from vegetation to the ground. This can be observed in Figure 4.11 between the vine rows where the transition from vine rows to ground level is slanted in the vertical profile. It can then be postulated that if the reference data for the absolute accuracy assessment had been collected at top-of-vegetation, the TPT accuracy over agricultural land would produce different findings reducing confidence in the current result.

A profile taken over an agricultural catchment dam, shown in Figure 4.12, revealed artefacts created by SfM over a body of water. In the SfM surface, there appeared to be a horizontal and vertical offset present on the dam wall (Left side of Figure 4.12).

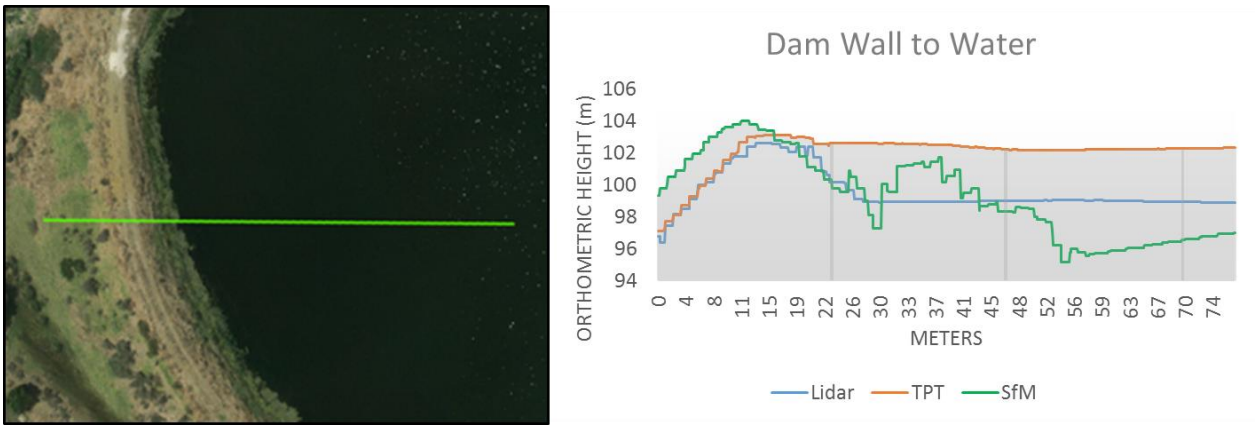


Figure 4.12 Transect of dam extending over the water body with image reference

Both TPT and LiDAR correctly reconstructed the water surface as a flat, horizontal feature visible to the right of Figure 4.12. SfM, however, failed to adequately model water as a result of variable contrast and the misidentification of key points of the non-static water surface. As the view direction and angle changes, specularly reflected features can result in the misidentification of key points in a SfM point cloud. Ripples on the water's surface coupled with sun reflection can create variable features. If these high-contrast features are visible in the imagery, this can produce incorrectly matched features across images generating inappropriate points with incorrect surface information. With no correct information to evaluate the epipolar geometry, these discontinuities and artefacts are created and represented in the final DSM. Specular reflection and weakly defined features are a known issue with SfM (Jancosek & Pajdla 2011). A possible solution would be to mask these areas out and interpolate a flat surface from the dam edges.

Transect profiles in built environments were purposefully chosen to include buildings of different sizes and densities across industrial, residential and informal settlement areas visible in Figure 4.13, Figure 4.14 and Figure 4.15 respectively. In these built environments, proximity and object size were the biggest contributors of error in image based reconstruction with little observed effect on LiDAR.

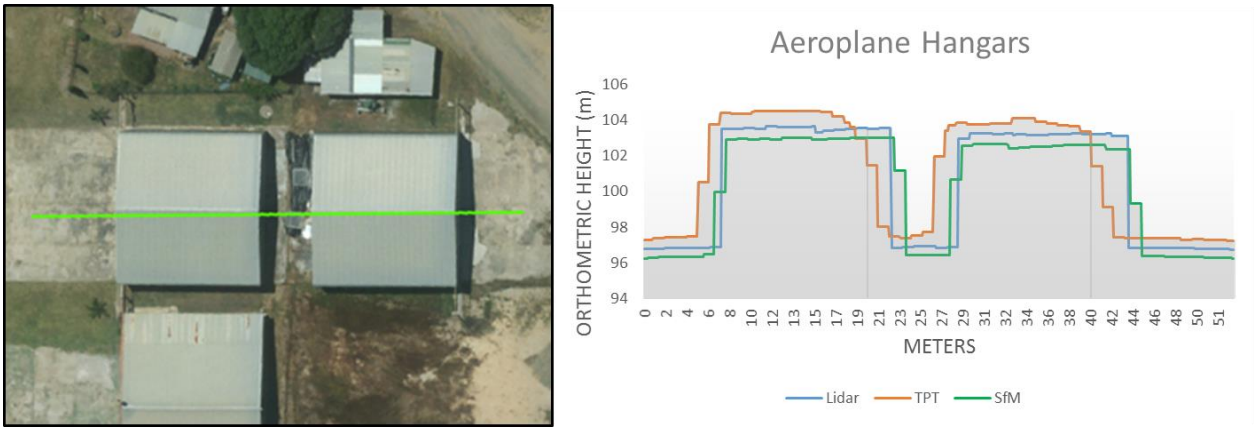


Figure 4.13 Transect of aeroplane hangars with image reference

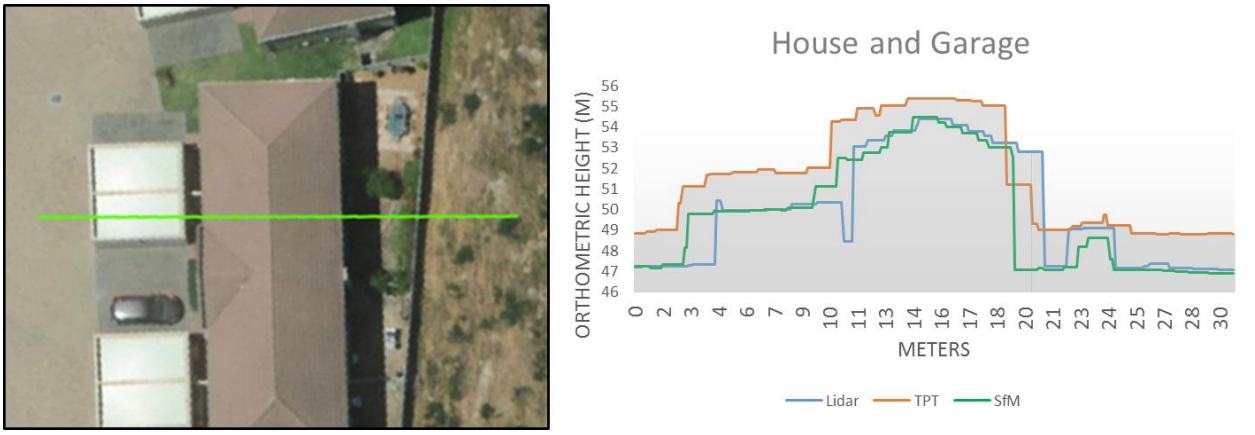


Figure 4.14 Transect of a house and adjoining garage with image reference

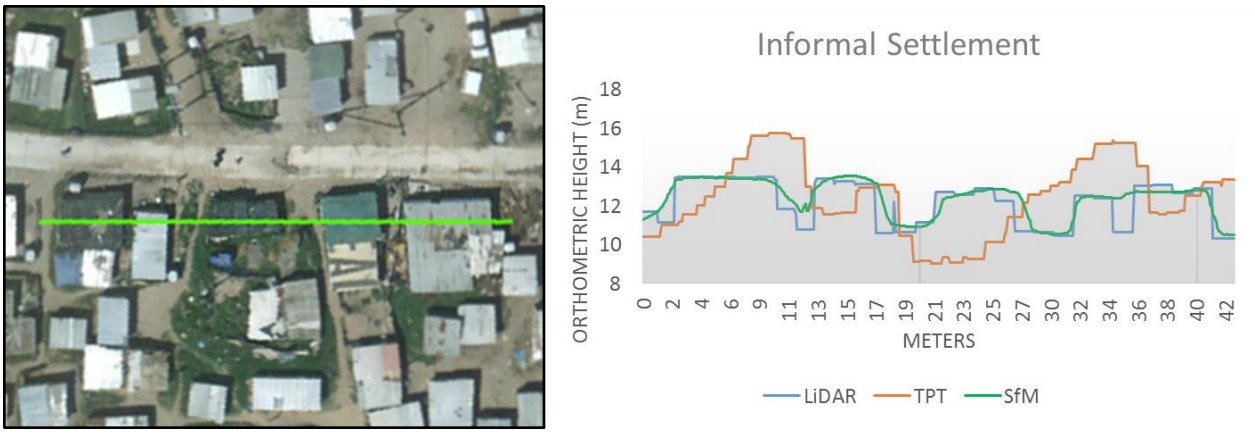


Figure 4.15 Transect of informal settlement dwellings with image reference

The larger plane hangars (Figure 4.13) were recovered adequately by all methods with LiDAR presenting the most accurate representation of the features structure. The rounding of corners was observed by TPT and tended to generalise the edges of this feature. SfM did not express this artefact to the same degree but extended the outer walls approximately 4-6 meters from the ground, which may be the result of point cloud surface meshing (Figure 4.13).

Despite flight path restrictions over urban areas, SfM managed to recreate the residential buildings, but both image-based DSM techniques (SfM and TPT) failed to separate the house and the garage (Figure 4.14). As a result of the purely oblique nature of the SfM image dataset, information to the west of the profile started to deteriorate due to object occlusion (Figure 4.7). The east side of the triangular roof was adequately represented, but detail increasingly deteriorated as image information was reduced on the west side of the roof. This also contributed to the over generalized surface connecting the house roof and the garage roof (Figure 4.14).

Over the smaller, highly clustered dwellings in informal settlements, it was found that TPT was incapable of correctly identifying the objects in the scene (Figure 4.15). An intrinsic problem in reconstructing informal settlements from imagery is that dwellings are typically constructed from corrugated galvanised steel. These roof structures are either painted or at varying levels of oxidation. A newer roof can be highly reflective and can cause lens flares in multiple directions, with large variabilities in brightness values and contrast. Similar to water bodies, specular reflection can severely limit the ability of SfM to reconstruct highly reflective objects. SfM did, however, manage to reconstruct vertical profiles of individual houses but less accurately than LiDAR. The surface estimation of SfM is generalized as a rounding effect that is observed on the edges of the wall structures. A failure to separate the buildings towards the east of the profile may be a limitation as close proximity of these dwellings is not uncommon.

4.2.3 Elevation variability over land cover

LiDAR is considered the most reliable technique for DSM construction as shown by previous studies (e.g. Stal et al. 2012; Johnson et al. 2014) and the total absolute error assessment (Section 4.1). It is therefore possible to use LiDAR as a reference to explore further the differences between SfM and TPT datasets in individual land cover classes. A GIS practitioner cannot assume that a single accuracy translates across all land cover types. Work performed by previous authors has produced accuracies over different land use areas and land forms but little information exists with

regards to the explicit accuracies over specific land covers. This section presents the SfM and TPT results using the statistical analysis of information extracted from subtraction of DSMs for each technique from LiDAR (simple subtraction raster) for different land covers.

4.2.3.1 LiDAR and SfM

Table 4.1 shows the aggregated variability in elevation across different land cover classes for SfM with respect to LiDAR from the simple difference rasters. The table shows that SfM DSMs best resembled LiDAR in vineyards given by the standard deviation of 0.902 m and the mean offset of the data mean of 0.386 m. This class also demonstrated the lowest range of data indicating fewer outliers within the observed DSMs across this land cover. The minimum and maximum values measured the largest negative and positive difference, respectively, recorded by a pixel. This could be a result of a spurious pixel value but is an indication of the range of these vertical artefacts.

Table 4.1 SfM-LiDAR simple difference raster with the calculated statistical variability in elevation over individual land cover classes

Land Cover Class	Min (m)	Max (m)	Range (m)	Mean (m)	Std. Dev (m)
Bare Field	-19.86	18.32	38.18	-0.11	1.01
Trees & Shrubs	-15.61	15.35	30.95	0.08	1.42
Low Vegetation	-19.55	8.07	27.62	0.06	1.02
Hard Surfaces	-20.55	18.15	38.69	0.12	1.32
High Density Residential	-39.90	5.22	45.12	0.05	1.04
Medium Density Residential	-15.86	17.61	33.46	0.33	2.39
Industrial	-17.91	19.30	37.21	0.58	3.56
Bare Ground	-21.70	17.69	39.39	-0.11	1.37
Vineyards	-2.83	4.59	7.43	0.39	0.90

Low vegetation also showed a good correlation to the LiDAR surfaces but trees and shrubs showed high data variability as indicated by the standard deviation of 1.418 m despite a low mean error of 0.082 m. This may be a result of temporal variation or could be an indicator of a horizontal offset since any XY displacement will result in high variability within the data (given by the standard deviation) and a low mean as a result of the averaging of this error. This is particularly true for tall

objects as they would have a greater impact on standard deviation in the presence of horizontal offset. Figure 4.16 illustrates a tree line from a simple difference raster showing vertical extrema as a result of this displacement. The extremely low values to the left of the trees (indicated by blue) and high values on the right (indicated by red) indicate the presence of a horizontal offset, the source of induced error. However, the proportion of error (area of high versus low values) is not purely from the offset and only accounted for some of the observed error. The clustering of low vegetation created over generalized objects in SfM reconstruction as SfM was unable to separate trees close together to the same degree as LiDAR (Figure 4.16).

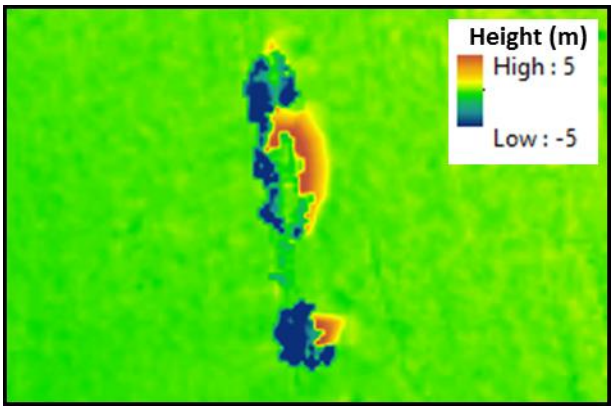


Figure 4.16 SfM-LiDAR simple difference raster over a tree line with difference height values given in meters

Industrial (density referring to structures per 100 m^2) areas were the worst represented by SfM with a high error range of 37.210 m and mean of 0.575 m and standard deviation of 3.556 m. The presence of offset was also observed in built areas, especially with larger buildings like that of industrial areas. Any offset on these larger buildings had a larger effect on standard deviation as it measures a squared relationship with distance away from the mean. This is in contradiction to the observed absolute accuracy (Section 4.1) where SfM demonstrated the lowest error of 0.282 m away from the GPS reference. The absolute accuracy assessment does not take into account the full quality of the DSM. The large homogenous roof structures on these industrial buildings were not reconstructed as geometrically rigid surfaces. These linear and low textured environments are a known limitation of SfM in surface generation (Jancosek & Pajdla 2011).

High density residential buildings (informal dwellings) demonstrated the lowest mean error of 0.049 m but the highest outlier range of 45.116 m. The proximity of dwellings to each other limited visual information captured by SfM imagery as a point needed to exist in three images to be

established. This limited SfM's ability to separate individual buildings and smoothed surfaces from the roof tops to the ground and was also observed by the cross-profile (Figure 4.15). A relationship was observed in built environments with an increased standard deviation and mean with increased object size and clustering.

Land cover classes such as bare ground and hard surfaces were not as well represented as anticipated by the low vertical variability with low/medium texture. In the instances of bare ground, a consideration is the spatial resolution of the classification. The pixel based classification of this land cover is 2.5 m spatial resolution but 10 m spectral resolution and may be prone to inclusion/mixing of vegetation (Table 3.1). Numerous pockets of bare ground are located as small clearings amongst shrubs and trees. Even though hard surfaces do not exhibit the same limitation as bare ground the means are similar suggesting that hard surfaces contain less textural information than expected. The standard deviations for bare ground (1.107 m) and hard surfaces (1.317 m) suggest a high variability in the elevation data despite hard surfaces having more linear profiles and yielding similar deviations to that of high density residential (1.038 m), trees and shrubs (1.418 m) respectively.

Bare fields demonstrated more variable terrain and are subject to low levels of regrowth due to temporal variation in the data capture. Notwithstanding, bare fields performed better that hard surface in a lower range of 38.132 m, mean -0.106 m and standard deviation of 1.006 m. This suggests that SfM is well suited to the surface reconstruction of agricultural land given that SfM also performed well in vineyards due to the more formal structure of vegetation growth and high textural information.

4.2.3.2 LiDAR and Stereophotogrammetry

Presented in Table 4.2 is the vertical variability over different land cover classes computed from the TPT-LiDAR simple difference raster. Bare fields performed well with low variability within the data given by the standard deviation of 1.027 m, irrespective of a relatively high mean of 0.740 m. This indicated that TPT is adequate for gradually sloping terrain and larger generalized land forms. Smaller and more linear classes like bare ground and hard surfaces performed comparably to each other suggesting more consistency with a range of 38.036 - 40.601 m, mean range of 0.759 - 0.832 m and standard deviation range of 1.439 - 1.569 m respectively.

Table 4.2 TPT-LiDAR simple difference raster with the calculated statistical variability in elevation over individual land cover classes

Land Cover Class	Min (m)	Max (m)	Range (m)	Mean (m)	Std. Dev (m)
Bare Field	-20.58	20.02	40.60	0.74	1.03
Trees & Shrubs	-19.87	18.35	38.22	0.12	2.17
Low Vegetation	-24.80	17.58	42.35	0.49	1.48
Hard Surfaces	-18.52	21.76	40.28	0.83	1.57
High Density Residential	-38.11	10.89	49.00	-0.08	2.67
Medium Density Residential	-14.94	19.30	34.25	1.85	1.99
Industrial	-15.47	20.24	35.71	3.20	3.29
Bare Ground	-18.29	19.75	38.04	0.76	1.44
Vineyards	-24.86	13.39	38.24	-1.59	4.03

High density residential showed the lowest mean elevation error of -0.075 m indicating a good average agreement to LiDAR but the vertical standard deviation of 2.667 m and range 49.000 m suggested large variability within the data. This is also evident in the cross profile (Figure 4.15) indicating an inability to represent dwellings despite a low mean. The high standard deviation reduced the confidence of any given point existing around the mean and in this case, the dwellings. As estimation technique is a function of squared distribution, the further a sample is taken from the mean, the greater its contribution to variance.

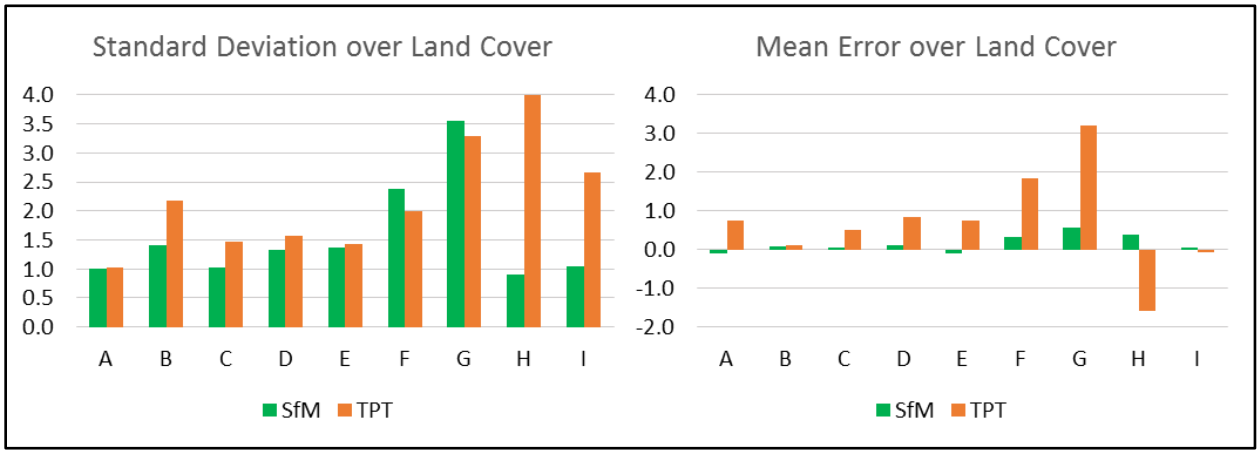
In medium density residential (i.e. suburban homes), TPT achieved better than undustrial and high density residential with a lesser degree of variability,1.987 m, and range, 1.852 m, in the built environment. The generalisation of large objects and the poor reconstruction ability of small objects suggested that TPT had increased error around the extremes of size and clustering. As with SfM, any horizontal offset in these rigid structures did result in high error variability calculations.

Vegetated classes showed no observable correlation with error range, mean and standard deviation. Taller vegetation, trees and shrubs, showed the least amount of mean error of 0.116 m, with a large distribution of over 2 m from the mean. Low vegetation performed better with a lower mean error from LiDAR and the lowest variation as a vegetative class of 1.475 m, suggesting it is capable of

adequately (depending on usage) handling low vegetation classes. Vineyards, with more formal structure, had the highest variation error of 4.032 m and a high mean error of -1.592 m indicating that TPT is inadequate for delineating vine rows which is congruent with the vertical profile (Figure 4.11).

4.2.3.3 Land cover elevation data exploration

Figure 4.17 illustrates the mean error and standard deviation of SfM and TPT against a LiDAR reference. The average mean error for SfM is 0.154 m compared to 0.703 m as the average mean error of TPT. This indicated that SfM has a better average vertical estimation to that of LiDAR. This measure does not, however, reflect variability of this average which is better reflected in the standard deviation for each land cover.



A	Bare Field	**F**	Medium Density Residential
B	Trees & Shrubs	**G**	Industrial
C	Low Vegetation	**H**	Vineyards
D	Hard Surfaces	**I**	High Density Residential
E	Bare Ground		

Figure 4.17 Standard deviation and Mean Error calculations per land cover class

The average standard deviation of SfM was 1.558 m compared with 2.184 m for TPT across all land cover classes. This suggests that SfM has less mean error and a higher probability of existing around that mean. SfM better simulated a LiDAR surface than TPT with a higher level of confidence that the surface mimicked that of the LiDAR surface. This statement applies in general terms and is in agreement with absolute error assessment and the cross-profile analysis. There was

however some differences across land cover types with higher variability in medium density residential and industrial buildings.

The Ecological Fallacy dictates that the average finding does not necessarily apply to the individual (Piantadosi et al. 1988), and, with respect to this assessment, not to the individual DSMs and individual land cover classes. For example an individual DSM may be more or less accurate than the total tabled average. Table 4.3 shows two distinct areas with some overlapping land cover classes. These are the differences between two SfM-LiDAR simple difference rasters showing a large variability in their results over common land covers for these individual areas.

Table 4.3 Individual overlapping land cover statistics for SfM DSM of an industrial area and farmland

Farmland					
Classification	**Min (m)**	**Max (m)**	**Range (m)**	**Mean (m)**	**Std. Dev**
Bare Field	-1.95	3.74	5.69	0.31	0.73
Trees & Shrubs	-1.12	1.72	2.85	0.09	0.68
Low Vegetation	-6.11	5.57	11.68	0.29	0.84
Hard Surfaces	-0.01	1.68	1.69	0.54	0.48
Industrial Area					
Bare Field	-19.86	17.42	37.28	0.18	1.64
Trees & Shrubs	-15.61	15.35	30.95	-0.030	2.78
Low Vegetation	-19.56	8.03	27.58	0.01	1.22
Hard Surfaces	-20.55	18.15	38.69	0.22	2.87

The range of error was notably higher in the industrial area (30.954 – 38.694 m) when compared to the farmland area (1.685 – 11.679 m). This shows that the outliers or extreme values were more prominent in this industrial area than in the farmland. Despite this, the mean error was less for each land cover class over the industrial raster compared to the farmland raster. However the higher standard deviation in the industrial raster indicates a lesser probability that the raster values actually existed close to that mean for every land cover. This means that the farmland area would better mimic that of the LiDAR. This was expected as these land cover classes, with the exception

of hard surfaces, are better represented in agricultural areas than in the industrial area. This result indicated that individual areas would express land cover accuracies that may not be representative of the overall findings for that land cover. This also suggests that an aggregated land cover assessment is a good indicator of expected ranges but may not be truly representative of individual DSMs.

4.2.4 DSM technique agreement over land cover

Isolating the lowest and highest values per unit area across LiDAR, TPT and SfM, a raster is created which indicates the areas with the largest amount of inter-model agreement and disagreement. These resulting extrema difference rasters were statistically evaluated across each land cover class to create Table 4.4 and gives an indication of DSM accuracy over different land covers. Agreement is measured by the mean error and the standard deviation. The lower the mean with a lower standard deviation, the higher the level of agreement each technique has with each other about the vertical reconstruction of that land cover.

Table 4.4 Land cover statistics computed from extrema difference rasters created with LiDAR, TPT and SfM combined

Land cover class	Min (m)	Max (m)	Range (m)	Mean (m)	Std. Dev. (m)
Bare Field	0.00	30.39	30.39	2.83	1.28
Trees & Shrubs	0.01	29.98	29.97	4.21	2.37
Low Vegetation	0.00	30.37	30.37	3.23	1.71
Hard Surfaces	0.00	30.45	30.45	3.52	2.06
High Density Residential	0.00	39.85	39.84	2.39	1.64
Medium Density Residential	0.01	19.28	19.26	2.82	2.23
Industrial	0.02	20.57	20.56	4.10	3.52
Bare Ground	0.01	30.50	30.50	3.46	1.68
Vineyards	0.00	25.83	25.82	2.91	2.67

The highest level of agreement was observed in bare fields and high density residential. Bare fields are large, gradually sloping areas that are high in texture lending themselves to high levels of DSM technique agreement with a mean error of 2.833 m and a high likelihood that a point exists near

the mean of 1.283 m. High density residential also indicated a high level of agreement with the mean error over these areas of 2.388 m with a low standard deviation of 1.643 m. This result appears anomalous given the variable accuracies previously calculated but allowing for a flat terrain and vertically short dwellings, a 2.4 m home is not atypical of the area, and with an allowance of 1.6 m above or below that height, the result becomes feasible. It does show the highest observed range of extreme values with outliers recorded over 9 m higher than the next highest range. As a collective mean, areas of higher agreement bring down the overall mean. In built environments, larger and taller structures exhibited increasing levels of disagreement shown by the increasing mean error and higher standard deviations. This suggested that technique choice becomes more of a factor with an increase in building size.

Flat and more linear land cover classes, hard surfaces and bare ground, showed higher mean error of 3.568 m and 3.455 m respectively with comparatively moderate variability in the dataset of 2.057 m and 1.676 m. Due to pockets of soil between vegetation and roads intersecting other land cover classes, DSM technique is an important consideration to accurately reconstruct these surfaces as they can contain spatially small features.

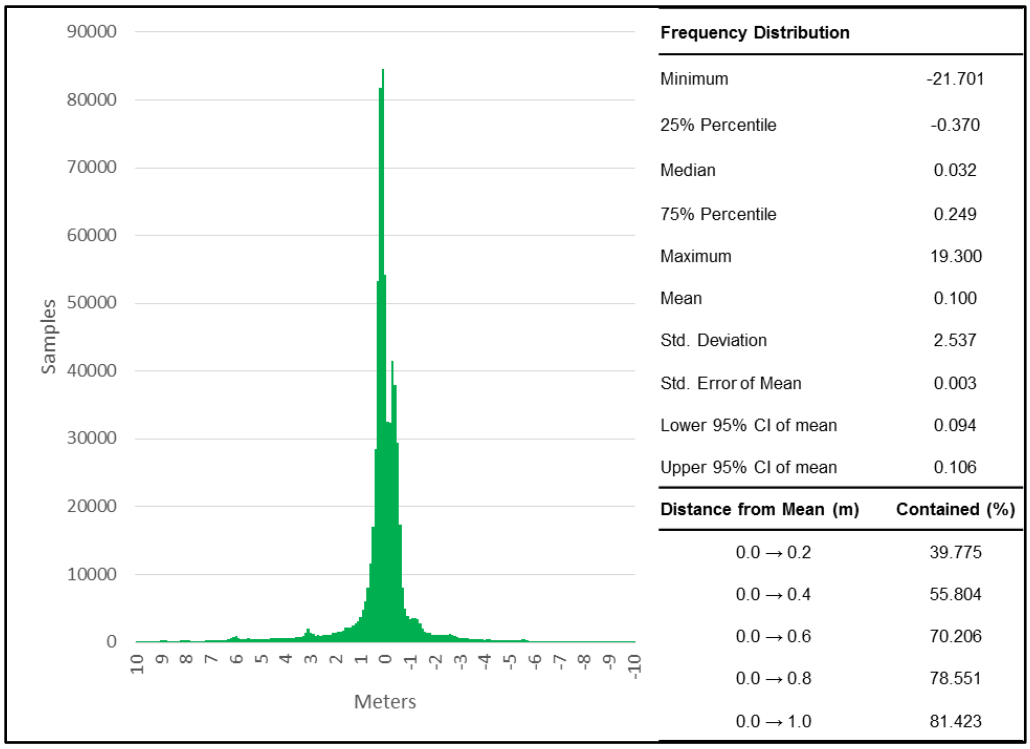
Vegetative environments illustrated some of the lowest levels of agreement and moderate to high variability within the dataset. Trees and shrubs showed the highest recorded level of disagreement with a mean error of 4.209m. Vertically tall but horizontally small features were not equally reconstructed in the DSMs with high levels of variability of 2.366 m. Low vegetation also showed a large mean discrepancy of 3.225 m but less variable data than that of trees and shrubs, shown by a low standard deviation of 1.709 m. Vineyards, although displaying the highest agreement as a vegetative class, showed low levels of agreement with mean error of 2.907 m and high standard deviation of 2.667 m. A study over vegetated areas with a dependence on vertical accuracy will need to review the technique choice as each technique will less likely agree with the reconstructed surface.

4.2.5 Uncertainty Analysis

The uncertainty or roughness analysis provides insight into the distribution of the vertical error (Clapuyt et al. 2015; Smith & Vericat 2015) of SfM and TPT from a LiDAR reference in different land use areas. The distribution of error in these areas is non-normal and does not satisfy the Central Limit Theorem (CLT) to assume normal distribution in non-normal datasets (Araujo &

Giné 1980). This means that standard deviation is an indication of variability but 68.27% will not be contained within the first deviation. As such, measures of 0.2 m intervals from the mean were presented with the data to describe the levels of inclusion or containment away from the mean that standard deviation would normally provide. By analysing the histogram, agreement in this context refers to inter-model correlation for specific land use classes from the simple difference rasters.

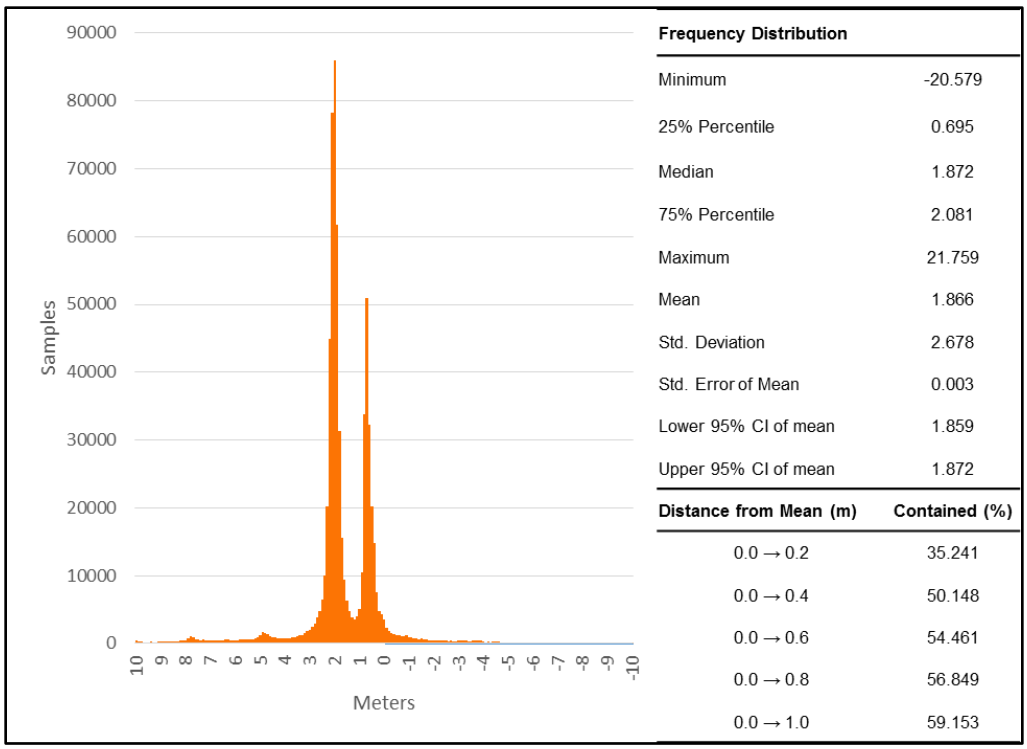
When comparing Industrial difference rasters (Figure 4.18 & Figure 4.19), the bimodal maxima's were identified and highlight inter-model disagreement which is more pronounced in the TPT-LiDAR rasters (Figure 1.19). Due to the sheer size of objects in industrial scenes, any horizontal offset will produce data spikes at intervals around or above 3 m from the mean; which has a compounding effect on the standard deviation, as error has a squared relationship.



Frequency Distribution	
Minimum	-21.701
25% Percentile	-0.370
Median	0.032
75% Percentile	0.249
Maximum	19.300
Mean	0.100
Std. Deviation	2.537
Std. Error of Mean	0.003
Lower 95% CI of mean	0.094
Upper 95% CI of mean	0.106

Distance from Mean (m)	Contained (%)
0.0 → 0.2	39.775
0.0 → 0.4	55.804
0.0 → 0.6	70.206
0.0 → 0.8	78.551
0.0 → 1.0	81.423

Figure 4.18 Data distribution of SfM-LiDAR difference raster in industrial areas



Frequency Distribution	
Minimum	-20.579
25% Percentile	0.695
Median	1.872
75% Percentile	2.081
Maximum	21.759
Mean	1.866
Std. Deviation	2.678
Std. Error of Mean	0.003
Lower 95% CI of mean	1.859
Upper 95% CI of mean	1.872

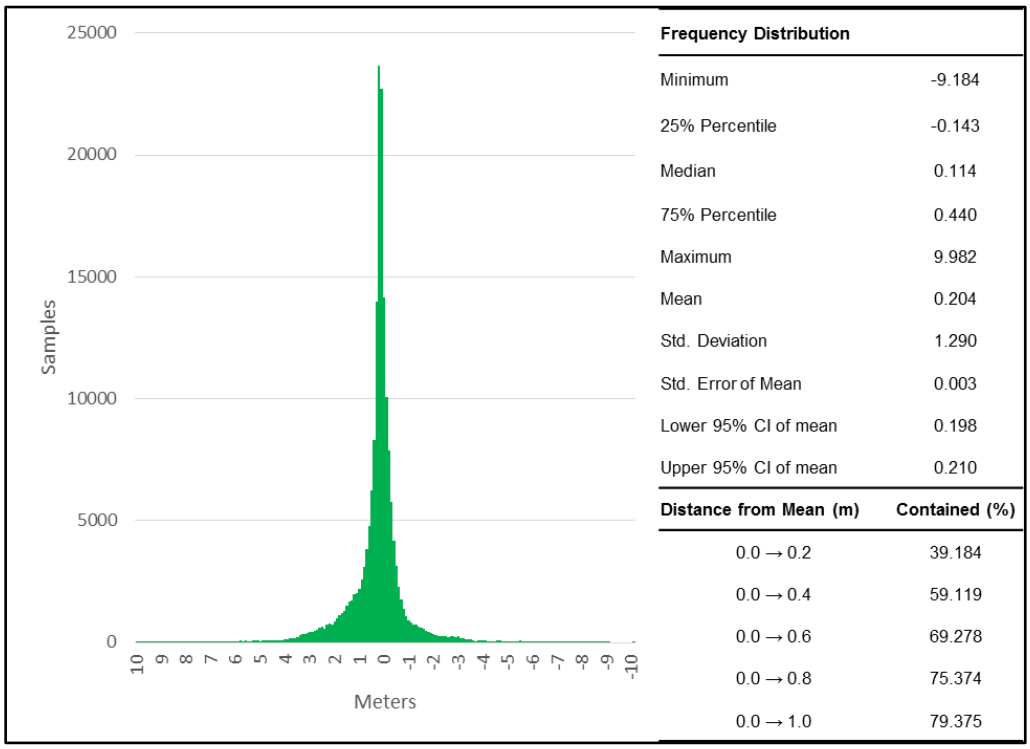
Distance from Mean (m)	Contained (%)
0.0 → 0.2	35.241
0.0 → 0.4	50.148
0.0 → 0.6	54.461
0.0 → 0.8	56.849
0.0 → 1.0	59.153

Figure 4.19 Data distribution of TPT-LiDAR difference raster in industrial areas

SfM showed that 81.4% of the data existed within a meter of the mean (Figure 4.18) whereas only 59.2% existed in the TPT data (Figure 4.19). The traditional photogrammetric data containment was measured from a mean that exists between two data peaks where the trough occupied a lower count of height samples. A sample of 50% of all elevation data was contained within 0.62 m from the mean whereas the same percentage of elevation data inclusion was observed 1.39 m from the mean in the TPT data. The standard deviation indicated there was a higher variance in the TPT dataset overall, confirming that SfM was more consistent and reliable in these areas.

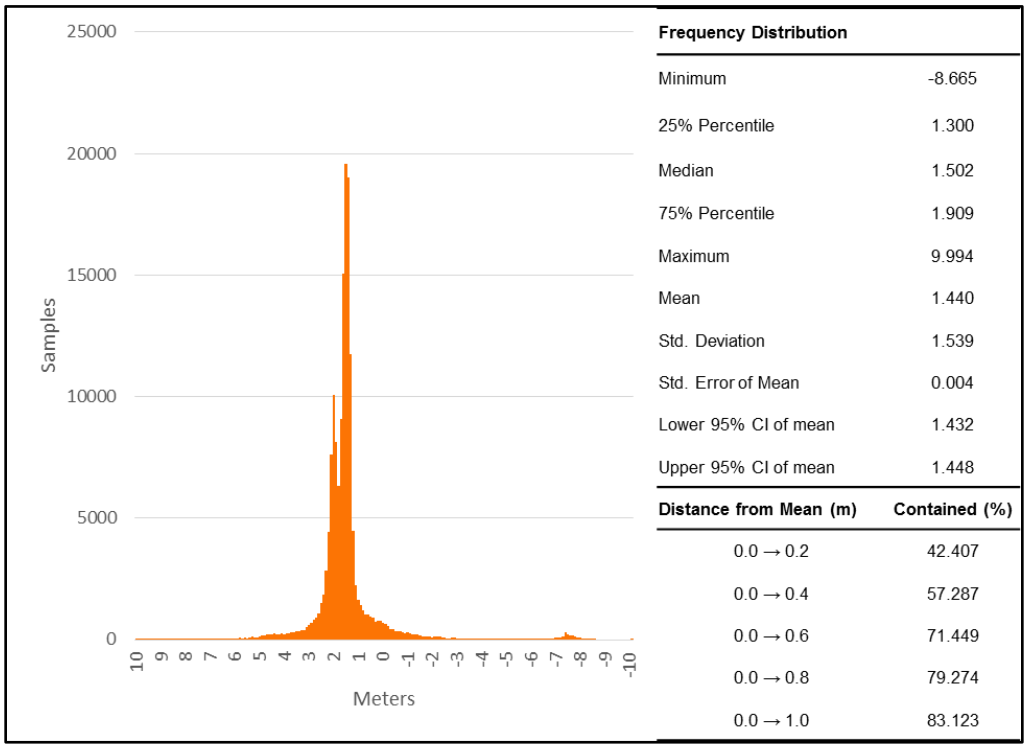
In residential areas the aggregated result from SfM illustrated a unimodal distribution with a single local maximum (Figure 4.20), which again suggested better inter-model agreement when compared to a bimodal TPT result (Figure 4.21). This variation of data distribution was also observed in the SfM simple difference raster given by a lower standard deviation of 1.290m when compared to TPT with a standard deviation of 1.539 m. Despite the lower standard deviation and higher inter-model agreement, a similar percentage of data from the mean was contained by TPT at 0.8 m compared to 1.0 m for SfM. At one meter from the mean there was only a 3.7% data containment deviation between the two methods. Given the scene and clustering of various objects

in a residential area, SfM and TPT performed comparably when compared to LiDAR based on their distributions.



Frequency Distribution	
Minimum	-9.184
25% Percentile	-0.143
Median	0.114
75% Percentile	0.440
Maximum	9.982
Mean	0.204
Std. Deviation	1.290
Std. Error of Mean	0.003
Lower 95% CI of mean	0.198
Upper 95% CI of mean	0.210
Distance from Mean (m)	**Contained (%)**
0.0 → 0.2	39.184
0.0 → 0.4	59.119
0.0 → 0.6	69.278
0.0 → 0.8	75.374
0.0 → 1.0	79.375

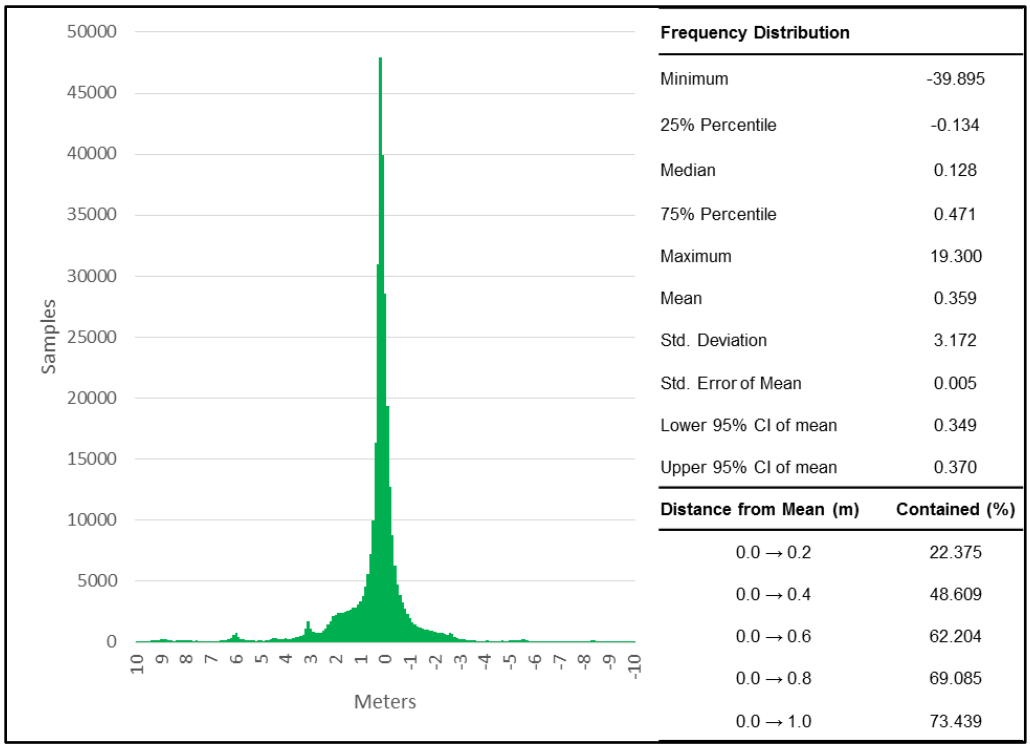
Figure 4.20 Data distribution of SfM-LiDAR difference raster in residential areas



Frequency Distribution	
Minimum	-8.665
25% Percentile	1.300
Median	1.502
75% Percentile	1.909
Maximum	9.994
Mean	1.440
Std. Deviation	1.539
Std. Error of Mean	0.004
Lower 95% CI of mean	1.432
Upper 95% CI of mean	1.448
Distance from Mean (m)	**Contained (%)**
0.0 → 0.2	42.407
0.0 → 0.4	57.287
0.0 → 0.6	71.449
0.0 → 0.8	79.274
0.0 → 1.0	83.123

Figure 4.21 Data distribution of TPT-LiDAR difference raster in residential areas

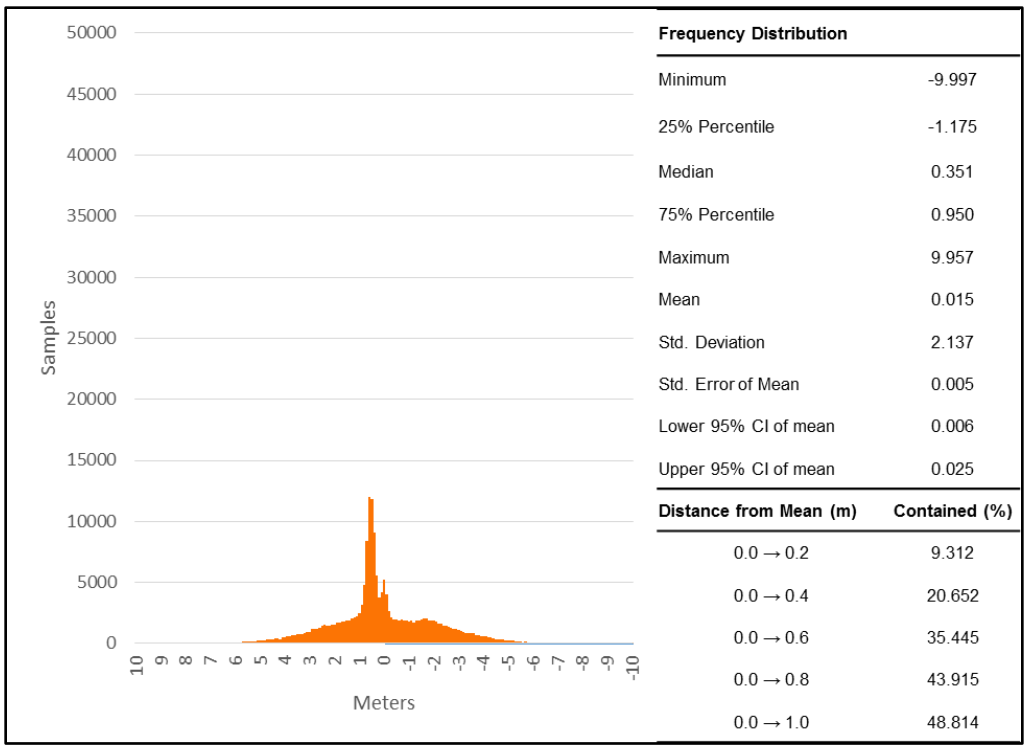
Figure 4.22 illustrates the data distributions of SfM-LiDAR data in informal settlements. 73.439% of the data distribution existed within a meter of the mean. Outlier peaks were observed at around the 6 m and 3 m intervals in the SfM histogram. Horizontal offset could be responsible for the peak (Figure 4.16) observed around 3 m due to the height of informal dwellings. Small peaks at 6 meters were likely due to the presence of a large number of power line structures not picked up by SfM but represented by LiDAR. TPT performed less well in informal settlements as seen from the absolute accuracy and the cross profiles results (Section 4.1 & 4.2.2). As illustrated in Figure 4.23, TPT achieved a low containment of 48.814% within a meter and high variability with a standard deviation of 2.137 m.



Frequency Distribution	
Minimum	-39.895
25% Percentile	-0.134
Median	0.128
75% Percentile	0.471
Maximum	19.300
Mean	0.359
Std. Deviation	3.172
Std. Error of Mean	0.005
Lower 95% CI of mean	0.349
Upper 95% CI of mean	0.370

Distance from Mean (m)	Contained (%)
0.0 → 0.2	22.375
0.0 → 0.4	48.609
0.0 → 0.6	62.204
0.0 → 0.8	69.085
0.0 → 1.0	73.439

Figure 4.22 Data distribution of SfM-LiDAR difference raster in informal settlements

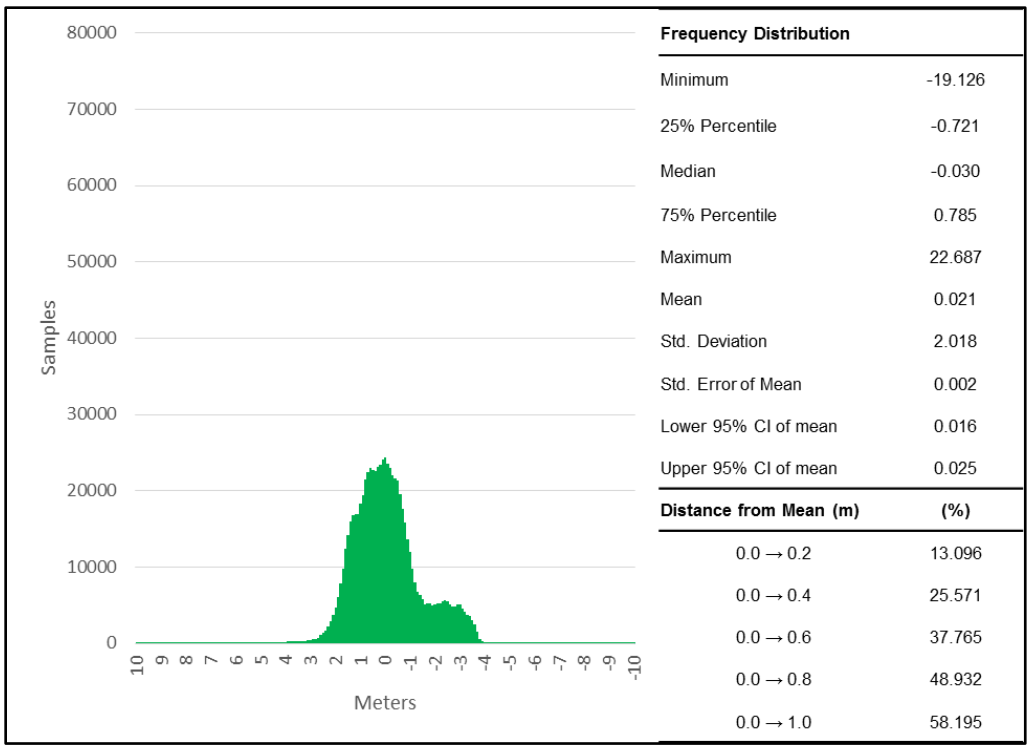


Frequency Distribution	
Minimum	-9.997
25% Percentile	-1.175
Median	0.351
75% Percentile	0.950
Maximum	9.957
Mean	0.015
Std. Deviation	2.137
Std. Error of Mean	0.005
Lower 95% CI of mean	0.006
Upper 95% CI of mean	0.025

Distance from Mean (m)	Contained (%)
0.0 → 0.2	9.312
0.0 → 0.4	20.652
0.0 → 0.6	35.445
0.0 → 0.8	43.915
0.0 → 1.0	48.814

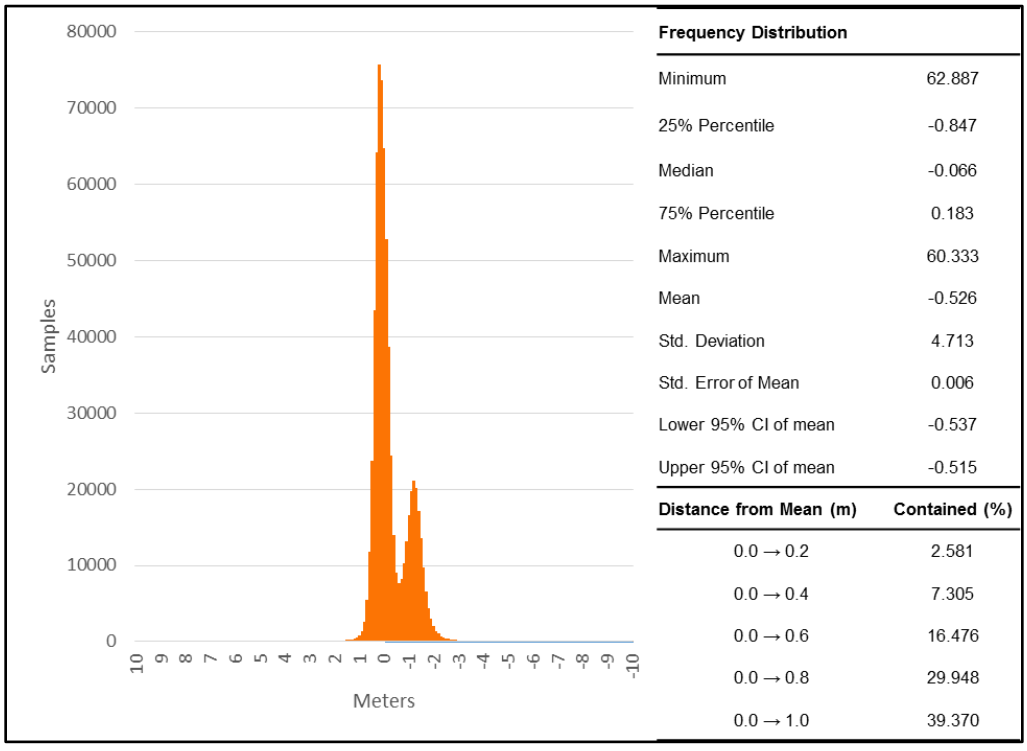
Figure 4.23 Data distribution of TPT-LiDAR difference raster in informal settlements

In agricultural areas both SfM (Figure 4.24) and TPT (Figure 4.25) performed relatively poorly with high levels of uncertainty within the data. SfM performed better, however, with a higher degree of inclusion and less variance. The containment of agricultural data with SfM was 58.2% inclusion 1 m from the mean, whereas TPT demonstrated 39.4% inclusion in the same area. Absolute error (Section 4.1) suggested a stronger correlation between SfM and LiDAR in the agricultural scenes and relative assessments (cross profiles and elevation variation over land cover) supported this finding. The bimodal histograms showed a high variation for both SfM and TPT with their distribution occupying a larger deviation.



Frequency Distribution	
Minimum	-19.126
25% Percentile	-0.721
Median	-0.030
75% Percentile	0.785
Maximum	22.687
Mean	0.021
Std. Deviation	2.018
Std. Error of Mean	0.002
Lower 95% CI of mean	0.016
Upper 95% CI of mean	0.025
Distance from Mean (m)	**(%)**
0.0 → 0.2	13.096
0.0 → 0.4	25.571
0.0 → 0.6	37.765
0.0 → 0.8	48.932
0.0 → 1.0	58.195

Figure 4.24 Data distribution of SfM-LiDAR difference raster in agricultural areas



Frequency Distribution	
Minimum	62.887
25% Percentile	-0.847
Median	-0.066
75% Percentile	0.183
Maximum	60.333
Mean	-0.526
Std. Deviation	4.713
Std. Error of Mean	0.006
Lower 95% CI of mean	-0.537
Upper 95% CI of mean	-0.515
Distance from Mean (m)	**Contained (%)**
0.0 → 0.2	2.581
0.0 → 0.4	7.305
0.0 → 0.6	16.476
0.0 → 0.8	29.948
0.0 → 1.0	39.370

Figure 4.25 Data distribution of TPT-LiDAR difference raster in agricultural areas

The overall distribution of the data suggested that SfM better emulated LiDAR in built environments with 73.4-81.4% data within one meter from the LiDAR surface compared to 48.8-83.1% from TPT. This suggests that SfM was more consistent when reconstructing these scenes. SfM, however, illustrated a larger standard deviation range (1.290-3.172 m) when compared to TPT (1.539-2.678 m) indicating that the DSMs may have a larger proportion of artefacts or outliers. The unimodal or close proximity of peaks in the bimodal profiles of the SfM histograms showed a superior inter-model agreement in surface reconstruction compared to TPT profiles providing further evidence of better consistency between SfM models. SfM demonstrated a larger inclusion percentage (58.2%) than TPT (39.4%) from LiDAR with both exhibiting poor inclusion results. The TPT bimodal histogram profile showed low levels of vertical agreement between models. The flat histogram profile of SfM suggests that it is not ideally suited for the reconstruction of vegetative scenes or the temporal offset regarding vegetative growth/harvesting of land distributed the error more uniformly.

4.3 ACCURACY ASSESSMENT SUMMARY

The accuracy assessments presented take two contextual forms: Land Cover and Land Use. The findings cannot be directly compared quantitatively but they were related to the nature and success of DSM construction. The focus on SfM was to better determine the appropriate applications of this technology in practice compared to the existing techniques of LiDAR and traditional photogrammetry. LiDAR was determined to be the most accurate reconstruction method followed by SfM and then TPT. SfM provided a higher level of detail than its traditional photogrammetric counterpart with more consistency and higher accuracies.

Success in built environments was determined by the nature of individual structures and their spatial proximity with other buildings and objects. SfM showed a relationship with increasing mean error and increasing standard deviation with building size and clustering observed in the land cover assessment when compared to LiDAR (Section 4.2.3.1). High density residential (informal settlements) exhibited overgeneralization around edges and caused noticeable variations within the dataset. Smoothing of these dwellings in SfM can become troublesome when delineating structures as sharp changes in height are more difficult to detect. In comparison, TPT showed little proficiency in reconstructing these scenes, evident from the cross profile (Figure 4.15), failing to identify individual dwellings. The overgeneralization of SfM and failure of TPT may a be a result of the spatial resolution, specular reflection of galvanised roofs, the clustering of the dwellings

themselves and the sun illumination providing low texture between structures. No single cause of this overgeneralization could be determined by this study but it is believed to be a combination of all these factors.

Vegetation is a constant limitation of image-based surface modelling (Johnson et al. 2014; Javernick et al. 2014; Tonkin et al. 2014) and careful consideration is necessary for localised studies. In agricultural crops SfM was proficient in defining individual crop rows which further substantiates its use in precision agriculture (Turner et al. 2012) where better image datasets are utilised. The vineyards land cover class showed low mean error and standard deviation when compared to land cover with cross-profiles showing adequate reconstruction of structure when compared to LiDAR. This is in contradiction to the absolute accuracy assessment but reference points taken on the edge of vine rows were determined to be the cause for this error for both LiDAR and SfM. Stereophotogrammetry and its propensity for the overgeneralisation of surfaces may present as a strength in terrain extraction rather than a surface extraction technique showing little propensity to recover structure from vegetative features. As a whole, vegetative areas in image-based techniques displayed the largest degrees of uncertainty with highly distributed data from the mean in all techniques.

From Westoby et al. (2013) we should observe 94% containment found within a meter or 90% of points values within 0.13- 0.41 m as determined by Johnson et al. (2014) for SfM data. The uncertainty analysis revealed the highest level of inclusion at 81.42% in industrial areas. This may be a result of the aggregation of multiple DSMs into one analysis instead of individual DSMs. A single industrial site measured 89.94% of data was contained within a meter of LiDAR (not the mean) which is within the range of Westoby et al. (2013) but was made lower with the inclusion of other DSM values. Uncertainty is exacerbated with a moving mean value and the creation of multimodal histograms will lower the level of inclusion. It has been stated that the conglomerate DSM values provide information pertaining to inter-model agreement and describes the consistency of error and the ability of data fusion which is not uncommon for large scale DEM constructions (Van Niekerk 2012).

LiDAR was determined the best technique as a whole for DSM constructions, retrieving accurate structural information with a high level accuracy. SfM was found better than the traditional photogrammetric technique recovering more structure, more accurate in absolute terms, better agreement of land cover classes to that of LiDAR with the least amount of uncertainty as a whole.

This however does not make TPT obsolete as it can process large areas with fewer images compared to SfM given the same camera and flying height. Legacy data is also captured for TPT compatibility and generally does not satisfy the overlap requirements of SfM. Both SfM and TPT are not competing techniques, but rather SfM is an extension of TPT with different uses determined by the size of the study site and the detail requirements.

4.4 FEASIBILITY CONSIDERATIONS

The various accuracies associated with LiDAR, stereophotogrammetry and SfM are an integral part of the decision making process to allow appropriate choice between these techniques but there exist internal and external factors that need to be considered and catered for within this decision. Practical deliberations are needed before the respective data can be captured or collected by the most appropriate technique to the purpose

Table 4.5 summarises the positive and negative factors affecting the feasibility of deploying DSM extraction techniques. The considerations include, but are not limited to, data sources (secondary data), data collection (primary data), sampling and resolution, expertise required, platforms, cost, processing requirements and repetition.

Table 4.5 Table summarizing the pros and cons of multiple elevation surveying techniques

	Ground Survey	**UAV SfM**	**Traditional Orthophotos**	**Ground Based LiDAR**	**Aerial LiDAR**
Type of elevation info	Terrestrial Point measurements	RGB stereo images	RGB stereo images	SpotLight Mode	IR light pulses blue or IR light pulses
Resolution	transects 50m apart with 10m point spacing	2 - 5 cm pixels	50cm pixels	1.5m pixels	3 points per m^2
Source of data	Total Survey Stations or Theodolite/ DGPS owners	Commercial Providers	NGI	Eg. CSIR	Commercial Providers
Sample density per km^2 for coastal application	2000 points	100 mil pixels	4 million pixels	Up to 200000 points per 5x5 mm	9 million points
Time to sample $1km^2$	1 day	10min (+ 4 hour field setup)	1.2 hours	2 hours	Seconds to minutes

Persons needed to sample	1-2	2	None	2	None
Approx. Cost per 1km^2	R8000	R5000	R50	R14300	R2000-R5000
Post processing required	Optional	Yes	Optional: GCPs	Yes	Yes
Can project manager do post processing	Yes	No	no, expert required	no, expert required	No, expert required
Cost of post processing	R1000 (2 hours)	R5000 (4 Hours)	R10 per 10km^2	R1500(fast)/day R500(slow)/day	-
Expected relative accuracy	Millimeters to centimeters	0.2-0.5m	0.5-2.0m	4mm horizontal x100mm vertical	Better than 10cm horizontal and vertical
Easy to repeat	Yes	Yes, for expert	Yes	Yes , for expert if benchmarked and surveyed	Yes, for expert
Data repetition	Anytime	Anytime	3 years	Anytime	Anytime
Typical application	General surveying, profiles, change detection	Local high-resolution surveying	Detect changes >1m, Land cover mapping	Monitoring erosion/accretion, Settlement and movement of features	high accuracy topography-related
Advantages of Technology	Easy, uncomplicated, doable by manager	Quick, cheap, high resolution	Cheap, not need for field work, high resolution	Relatively fast, not wind sensitive, can be operated by 1/2 man team	fast, covers large areas, high accuracy, established technology
Limitations of technology	usually low point density, i.e. little detail, or very time consuming	sensitive to wind; not for large scale coverage; legal issues on flight permits; technology	sophisticated post-processing: artefacts in water and low-contrast areas; low image repetition	Shadows/gaps as a result of survey platform being close to ground; site access	Weather Dependent

Adapted from Luck-Vogel (2015, Pers com)

Human capital is a consideration that has to be deliberated when processing data. SfM is, to a large extent, automated (James & Robertson 2012) but this advantage is not specific to SfM alone. Aerial imagery for stereophotogrammetry from vendors is typically provided with an RPC file which can demonstrate a similar level of automation as SfM would with geotagged imagery. This level of georeferencing imagery can be less accurate and the spatial error will be reflected in the final DSM. Considerations of this nature also extend to LiDAR only requiring the appropriate return filtering and interpolation. The SfM workflow is easy to process not requiring expert knowledge but some basic knowledge is needed to operate the technology. This however is offset by the need for a trained and certified pilot – a limitation reflected in all primary data capture from an aerial platform.

UAVs can offset costs, but flight restrictions and the areas that can be sensed are smaller. Lightweight UAV LiDARs are in production but are more costly than commercial cameras and significantly heavier. Compensation for the weight can be solved with high powered platforms commonly associated with multirotors. These aircraft are prone to shorter flight durations of 25 minutes or less and cover tens of acres (10ac ≈ 4ha), where fixed-wing platforms can cover hundreds of acres (100ac ≈ 40ha) (Van der Merwe & Price 2015). If SfM is mobilised with the use of UAV (RPAS) platforms, legislative flight restrictions apply and appropriate permission may not be guaranteed by the director of the SACAA with the current strict regulation governing populated areas (South African Civil Aviation Authority 2015). SfM and LiDAR are not mutually exclusive techniques for DSM creation. The point clouds can be merged together in an attempt to capitalise on the strengths of both techniques assuming a relatively static scene (Meesuk et al. 2015). SfM can provide high levels of detail and edge information while the accuracy of LiDAR's laser ranging can still preserve accurate structural information.

In the instance of secondary data collection, variability is prominent in the data with regard to availability and spatial resolution. Governing bodies and NGOs collect data both for stereo photogrammetry and LiDAR. In the Western Cape aerial imagery is captured for the whole province by NGI but CoCT captures images and LiDAR for the Cape Town metropolitan area only. These datasets typically differ in spatial resolution dependent on their intended purpose by the individual collection agencies. Additionally the data is collected in 1-3 year intervals producing temporal limitations (Table 4.5). The aerial imagery collected as yet does not contain adequate overlap for SfM based processing. This would require the combination of temporal datasets (data

taken at different dates) or manual imagery capture processes. For TPT and LiDAR the scope of study would have to fit within the time frame of capture unless the observed features/objects are temporally static (i.e. buildings or large rock formations). LiDAR, in this particular study, was confined to the Cape Town metro requiring the study site to exist within the available extent. The situation in South Africa is not true in all situations since in some countries, such as the Netherlands, national LiDAR surveys have been completed and are openly available (Actueel Hoogtebestand Nederland 2015). If vendors change their flight patterns to accommodate SfM, overheads can be expected to increase with the larger datasets and more compact flight patterns. Even if this is employed as a routine part of commercial image capture, TPT will still be needed for any legacy processing and for areas larger than 8000m2 (Siebert & Teizer 2014). Currently large areas with a high resolution are impractical for SfM computation due to the exponential time increment in processing with every added image. The size of the output point clouds and the level of detail needed is a consideration. Determining a digital surface of a provincial extent, individual crop rows may be unnecessary for a DSM project at that level making TPT a better choice in this instance.

Imagery as a source of information has potential for more than just DSM creation as it intrinsically contains spectral data. A single dataset can produce topographical and spectral information contained within a scene. With multispectral sensors a large number of spectral indices can be computed in conjunction with surface elevation, providing a vehicle for aggregated information gathering with no temporal offset. The accuracy of the spectral and spatial accuracy is dependent on the capabilities of the camera and the flying height. The predicted accuracy of SfM as a whole can be difficult to estimate owing to a multitude of variables: spatial resolution; overlap; capture angle; scene complexity; scene texture; clustering of objects; coverage; sun angle, object illumination etc.. SfM-based modelling is a function where these factors contribute to the quality of the reconstructed scene. Research for commercial applications will typically take the form of airborne platforms to reduce issues intrinsic to terrestrial techniques (i.e. dead ground).

Although the fundamental algorithms that produce a SfM DSM are known, they are also subject to proprietary alterations, optimisations and algorithmic substitutions to achieve similar surface reconstructions. Specific information, however, about the true nature of this processing may not be directly in the public domain. This can lead to the ‘black box’ perception (Remondino et al. 2014; Micheletti et al. 2015) of processes in proprietary software which can limit scientific

exploration in application-based research. The customisation of the SfM parameters can vary between software packages and with varying skill levels. A possible repercussion may include varying results from techniques across different software packages. For example, Designing Reality puts an emphasis on catering for weak features handling homogenous areas and reflective surfaces, potentially producing a better output based on the research of Jancosek & Pajdla (2011).

CHAPTER 5: CONCLUSION AND RECOMMENDATIONS

This chapter summarises the key findings from the results of the six accuracy assessments and integrates these with the practical considerations of utilising SfM. The limitations of this study and recommendations for future research to extend and improve the knowledge base around SfM are presented as well as concluding remarks.

5.1 SUMMARY OF FINDINGS

The aim of this study was to compare SfM to well understood commercial techniques (traditional stereophotogrammetry and LiDAR) in DSM creation in varying environments. This was achieved by means of three assessments: (i) absolute accuracy, (ii) relative accuracy and (iii) the limitations of these accuracy assessments coupled with feasibility considerations.

(i) Absolute accuracy assessment

It was found that LiDAR was the best method for surface reconstruction picking up the highest level of detail and overall the lowest margin of error over different land uses (RMSE 0.57m). SfM in the same assessment found an overall RMSE of 0.67 m with stereophotogrammetry performing least well with an RMSE of 1.22 m. An average of 0.34 m offset was observed between LiDAR and SfM from an independent GPS reference which is congruent with literature (Section 2.4). In built environments, SfM was compared to an independent GPS reference, and demonstrated good results with an RMSE range of 028-0.62 m. A large vertical error in agricultural environments was consistently observed with range of 0.76-1.7 m across all agricultural sites. In these areas, the SfM-LiDAR relationship demonstrated a 0.08 m error disparity margin confirming SfM's comparability with LiDAR (Westoby et al. 2013; Fonstad et al. 2012). As demonstrated by the high vertical error. TPT performed poorly overall with a 1.23 m RMSE compared to 0.67 m measured by SfM.

(ii) Relative accuracy assessment

Variability in land cover type is a confounding factor in the accuracy of all forms of DSM surface reconstruction. SfM failed to separate objects in close proximity which was identified as a limitation both in informal settlements and non-formal vegetation. SfM reconstruction of vegetation, in particular, had negative effects on DSM accuracy. In a subtractive comparison to LiDAR, agricultural areas demonstrated 58.2% of the data was contained within a meter from the

mean despite transects illustrating similar surface structures. SfM did, however, pick up formal crop rows and demonstrated a low mean error from a LiDAR surface reference in the vineyard land cover class. Stereophotogrammetry failed to pick up vine rows and displayed higher levels of overgeneralization in vegetated areas with the highest levels of uncertainty.

SfM performed well in built environments but with decreasing accuracy with increased clustering. It was found that in built environments 73.4 – 81.4% of the data existed within a meter of the mean suggesting high correlation with more linear features. Success in the built environment was determined by the nature of individual structures and their spatial proximity with other buildings and objects. This finding was in contrast with the cross-profile analysis as smaller buildings showed less propensity to preserve shape. It was established that this was a result of horizontal offset as larger buildings can markedly affect the distribution of data, as observed in the uncertainty histogram.

The uncertainty analysis revealed the highest level of inclusion at 81.42% in industrial areas by SfM. This was lower than expected from the literature but may be a result of the aggregation of multiple DSMs into one analysis instead of individual SfM DSMs. A single industrial site measured 89.94% of data contained within a meter of LiDAR (not the mean) which is close to the range of Westoby et al. (2013) but was made lower with the inclusion of other DSM values. Uncertainty is exacerbated with a moving mean value in the data distribution histograms which lowers the level of inclusion. The creation of more multimodal histograms by TPT suggested that it was a less well suited technique for consistent DSM construction when compared to SfM.

(iii) Limitations and suitability of SfM

Land cover statistics confirm more linear areas and larger objects were more successfully reconstructed by SfM. Weaknesses were observed over water bodies, as observed in the cross profile, with failures in reconstruction when specularly reflective features hindered accurate point matching. The inter-DSM agreement of SFM in built environments suggested better propensity for data fusion than stereophotogrammetry. The ability to reconstruct built features was adequate but limited by highly clustered buildings resulting in overgeneralized surfaces which represents an additional weakness of SfM in areas with high density residential.

Traditional photogrammetric reconstruction performed least well in this study demonstrating the lowest accuracies and level of detail with the exception of absolute accuracy pertaining to

agricultural areas. This was found to be a result of GPS validation points taken close to crop rows where vegetative growth obstructed ground reconstruction of LiDAR and SfM. Traditional photogrammetry showed its best results in residential settings with a higher inclusion of data within a meter of the mean but illustrated a higher variability in the dataset which implies poor reproducibility.

SfM in relative performance demonstrated a high level of congruency in object identification. ALS is the preferred technique for DSM generation as it can covers large areas, is highly accurate and is a well-established technology. However this technology does require specialist knowledge. For local scenes in areas difficult to access, requiring highly temporal data, SfM is better suited to the task. Information can be derived on site without the need for specialist support. This enables researchers to capture data from remote regions and from multiple datasets to reconstruct an area for study.

In the Western Cape aerial imagery is captured for the whole province but high resolution imagery and LiDAR is only available for the Cape Town metropolitan area which necessitates primary data capture for studies planned outside the metropolitan area. Given a small study area (100-100 000 m^2) with limited skills and budget, SfM with UAV technology would be a suitable substitution for LiDAR (Siebert & Teizer 2014).The benefit of using imagery for primary data capture as it has potential for more than just DSM creation as it intrinsically contains spectral data.

The predicted accuracy of SfM as a whole can be difficult to estimate owing to a multitude of variables: spatial resolution, overlap, capture angle, scene complexity, scene texture, clustering of objects, coverage, sun angle, object illumination etc. SfM-based modelling is a function where these factors contribute to the quality of the reconstructed scene. This was in agreement with Westoby et al. (2012: 302) “it is impossible to offer explicit guidance on the minimum number of photographs necessary for successful scene reconstruction”. Specific information, however, about to the true nature of this processing may not be directly in the public domain. This can lead to the ‘black box’ perception (Remondino et al. 2014; Micheletti et al. 2015) of processes in proprietary software which can limit scientific exploration in application-based research.

The highest level of DSM inter-technique agreement was observed in bare fields and high density residential. Bare fields are large, gradually sloping areas that are high in texture lending themselves to high levels of DSM technique agreement with a mean error of 2.83 m and a high likelihood that a point exists near the mean of 1.28 m. Vegetative environments illustrated some of the lowest

levels of agreement and moderate to high variability within the dataset having recorded levels of disagreement with a mean error of 4.21 m. Vertically tall but horizontally small features were not equally reconstructed in the DSMs with high levels of variability of 2.37 m. Low vegetation also showed a large mean discrepancy of 3.23 m. Vineyards, although displaying the highest agreement as a vegetative class showed low levels of agreement with mean error of 2.91 m and high standard deviation of 2.67 m. A study over vegetated areas with a dependence on vertical accuracy will need to review the technique choice as each technique will less likely agree with the reconstructed surface.

5.2 CONTEXTUALISATION AND CONTRIBUTION

As discussed in the formulation of the study aims (Section 1.2), the broader uptake of SFM as an operational tool is hindered by the current progress of research and literature base. SfM needs to be assessed in terms of reliability, accuracy and precision in a wide range of environments before it can be considered as an accepted practice in GIS (Fonstad et al. 2012). A rigorous accuracy assessment is required over multiple land forms and land types to better elucidate the major sources of error, which is essential for the adoption of this method in practice (Westoby et al. 2012).

5.2.1 Adding to the body of knowledge

The contribution of this research begins to address the concerns, brought forward by Westoby et al. (2012) and Fonstad et al. (2012). This study evaluated the vertical accuracy of SfM in terms of both absolute and relative accuracy. The multiple relative assessments allow an appreciation of the nature of error and contribute an understanding to strengths and weaknesses in all three techniques pertinent to technique choice and study design for future use and research. The vertical accuracy, which predicts DSM quality and reproducibility in existing and accepted techniques, has shown that SfM compares favorably with existing techniques. This can, in conjunction with existing literature, aid in the acceptance of SfM in GIS and contribute to the data available to identify the most suitable techniques in differing land classes.

To echo the words of Green et al. (2014: 181) “Until structure from motion can demonstrate reliable accuracy, and this can be calculated on a case by case basis, it is unlikely to be taken as seriously as a measurement tool”. This research was performed over varying land forms and land types by assessing both land use and land cover types and has provided various accuracy assessments for each land cover type that adds to the existing body of knowledge. These

environments are posed as a case study as results vary over different environments even if the land cover/use classes are the same (Section 2.5). Although the results speak only to the accuracy over this area, the volatility of accuracy results are a key factor in the implementation of LiDAR, TPT and specifically SfM. The reliability of SfM, tested in this study, has shown proficiency in built environments with results having indicated higher levels of reproducible accuracy compared to stereophotogrammetry which is currently the accepted measurement tool for the majority of existing DSM's available in South Africa. Individual accuracies, however, may need to be calculated for each DSM as various individual accuracies have been shown to fluctuate in differing environments.

5.2.2 Limitations of study

As image-based DSM techniques, both SfM and TPT use independent image datasets. Despite the fact that the output DSMs were resampled to the same resolution, any differences in dataset (spatial resolution, lens distortion, sun angle etc.) are not accounted for in this study. This was unavoidable because of differing overlap requirements by the TPT imagery in conjunction with the availability of image data with appropriate platforms and vendors. The reconstructive ability of both SfM and stereophotogrammetry algorithms would have been better observed using a single dataset. This would also have served to eliminate any temporal offset that may have been present in the dataset.

The consistency of flying height and the amount of oblique imagery hindered the reproducibility of the results with different image datasets. This is not to say that they may not have been similar, but any variation (internal or external) within the imagery could result in different surface extents and disparity in the quality of the output DSM.

As this study used UAV-simulated data, actual UAV data could differ. Assuming the same spatial resolution and distortion, the field of view (FOV) may not have remained constant affecting the overlap between images or the number of images. An ancillary observation could look into the computation time of larger datasets with the same amount of overlap. The increase in overlapping imagery, coupled with a non-stable aerial platform, could produce superior DSM results as more information would be made available about the epipolar geometry of the observed surface.

5.3 RECOMMENDATIONS FOR FUTURE RESEARCH

From this study, further research questions have arisen to better understand the suitability of SfM deployment:

a) A study into the vertical accuracy and reconstruction ability of traditional photogrammetry and SfM using a single GPS and image datasets would provide more conclusive and accurate comparisons into the strengths of the mathematical algorithms used for DSM construction while limiting the input data as a source of error.

b) Pix4D Mapper is one of many software solutions in the commercial sector. The problem of 'black box' processing with proprietary algorithms has unknown effects on the final output accuracy. A review of the available software packages with controlled parameters and input data could quantify the inter-software agreement of DSM construction and predict advantages and disadvantages of the use of certain SfM software packages available.

c) Studies into field of view (FOV) with static overlap and spatial resolution would help to illustrate the breaking points of homogeneity in land cover types and possibly produce better guidelines for SfM data acquisition from aerial platforms. This study could be extended to describe DSM quality and vertical accuracy using a consistent overlap percentage, given the same spatial resolution and lens distortion, into the effects of FOV and dataset size in DSM creation.

d) As a single study, directly observing land cover and land use, the conclusions of this study would become more robust if supported by comparable or complementary research. This would further allow a better understanding of the overall performance of SfM in DSM reconstruction in varying land cover areas and lead to more informed predictions into the performance of SfM for future positioning as a research tool.

5.4 CONCLUSION

The generation of 3D surface constructs is an emerging and valuable tool in GIS and allows for multiple applications over all land types including the capturing of population based and land use information. This includes the need to acquire accurate documentation of dwellings and infrastructure for planning, maintenance and compliance, disaster assessment and relief, industrial land use and agricultural capacity including sequential crop data. Current techniques for the generation of DSM reconstruction models are led by LiDAR in terms of absolute accuracy and this is confirmed by this study. However SfM results were comparable and demonstrated superiority to traditional photogrammetric reconstruction in both accuracy and level of detail. This is in line with the existing data on SfM accuracy from a growing literature base. SfM reconstruction is dependent on a multitude of factors which may vary over different land use types and therefore expected accuracy must be expressed as a range with consideration given to the composition of the scene.

SfM has been found to be practical and comparable to LiDAR without requiring expert skill and is largely automated making it an effective tool in GIS for local and remote regions. There are however some practical considerations which currently hinder the mobilisation of SfM based studies which GIS practitioners should consider when deciding on the appropriate technology. These include legislative constraints governing the use of UAV platforms especially limiting their use in urban areas. Clearly this may change with evolving legislative adjustments and these can be informed by the improved appreciation of the applications of land use mapping within municipal and government planning.

SfM has better than the traditional photogrammetric technique recovering more structure, more accurate in absolute terms, better agreement of land cover classes to that of LiDAR with the least amount of uncertainty as a whole. Despite the inferior performance, traditional photogrammetric techniques will not be replaced on account of dataset size, existing legacy data and large area surveys. Instead we expect and are currently observing SfM principles being incorporated into stereophotogrammetric techniques as evidenced by the adoption of the SIFT algorithm in select software packages. SfM should rather be considered the natural extension of photogrammetric processing for use in smaller, highly detailed and temporally variable environments.

REFERENCES

Actueel Hoogtebestand Nederland 2015. AHN system [online]. Amersfoort: AHN. Available from http://www.ahn.nl/pagina/open-data.html [Accessed 5 October 2015]

Agarwal S, Furukawaa Y, Snavely N, Simon I, Curless B, Seitz SM & Szeliski R 2011. Building Rome in a day. *Communications of the ACM* 54(10): 105-112.

Ahmad A, Leh LC & Udin WS 2012. *Accuracy assessment of Aerial Triangulation using Different Format of Aerial Photograph and Digital Photogrammetric Software*. Proceedings of the IEEE 8th International Colloquium on Signal Processing and its Applications (CSPA) held 23-25 March 2012. Melaka, Malaysia.

Agugiaro G, Poli D & Remondino F 2012. Testfield Trento: Geometric Evaluation of Very High Resolution Satellite Imagery. International Archives of the Photogrammetry, *Remote Sensing and Spatial Information Sciences* 39: 191-196.

Araujo A & Giné E 1980. *The central limit theorem for real and Banach valued random variables*. New York: Wiley.

Albertz J 2007. A look back. *Photogrammetric Engineering & Remote Sensing* 73(5): 504-506.

Arya S, Mount DM, Netanyahu NS, Silverman R & Wu AY 1998. An optimal algorithm for approximate nearest neighbour searching in fixed dimensions. *Journal of the ACM* 45(6): 891-923.

Bater CW & Coops NC 2009. Evaluating error associated with Lidar-derived DEM interpolation. *Computers & Geoscience* 35: 289-300.

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

Bolles RC & Fischler MA 1987. Random sample consensus: a paradigm for model fitting with applications to image analysis and automated cartography. *Communications of the ACM* 24(6): 381-395.

Brown DC 1971. Close-range camera calibration. *Photogrammetric Engineering* 37(8):855-866

Brunn A & Weidner U 1997. Extracting Buildings from Digital Surface Models. *IAPRS* 38(3) 27-34.

Campbell JB 2006. *Introduction to Remote Sensing*. 4th ed. Abingdon: Taylor & Francis

Cheuk ML & Yaun M 2009. Assessing Spatial Uncertainty of Lidar-derived Building Model: A Case Study in Downtown Oklahoma City. *Photogrammetric Engineering & Remote Sensing* 74(12): 257-269

City of Cape Town SDI and GIS Department 2015. City statistics and population census [online]. Available from: https://www.capetown.gov.za/en/stats/Pages/default.aspx [Accessed 9 July 2015]

Chandler G & Merry C 2010. The South African geoid 2010: SAGEOID10. *Position IT*. June: 29-33.

Chen S, Wang Y & Cattani C 2012. *Key Issues in Modelling of complex 3D Structures from Video Sequences*. Mathematical Problem in Engineering 2012.

Clapuyt F, Vanacker V & Van Oost K 2015. Reproducibility of UAV-based earth topography reconstructions based on Structure-from-Motion algorithms. *Geomorphology* [online]. Available from: http://www.sciencedirect.com/science/article/pii/S0169555X15002652 [Accessed 9 July 2015]

Congalton RG & Green K 2008. *Assessing the accuracy of remotely sensed data: principles and practices*. 2nd ed. Boca Raton: CRC press.

Cornelis N, Leibe B, Cornelis K & Van Gool L 2008. 3D urban scene modeling integrating recognition and reconstruction. *International Journal of Computer Vision* 78(2-3): 121-141.

Crandal D, Owens A, Snavely N & Huttenlocher 2011. *Discrete-continuous optimization for large-scale structure from motion*. Proceedings of the computer vision and pattern recognition conference held on 20-25 June 2011. Providence: IEEE.

CUDA 2013. *CUDA Programming Guide*. NVidia. [Online Help] Online available: http://docs.nvidia.com/cuda/cuda-c-programming-guide/index.html.

Deems JS & Painter TH 2006. *Lidar measurement of snow depth: accuracy and error sources*. Proceedings of the international snow science workshop 16 Oct 2006. Telluride: ISSW

Dietrich JT 2014. Applications of Structure-from-motion Photogrammetry to fluvial geomorphology. Doctoral Dissertation, Department of Geography, Graduate School of the University of Oregon.

Ekhatari N, Sahebi MR, Valadan Zoej MJ & Mohammadzadeh A 2008. Automatic Building Detection from LiDAR Point Cloud Data. The International Archives of the Photogrammetry, *Remote Sensing and Spatial Information Sciences* 37: 473-478.

ESRI 2012. *ArcGIS Resources for ArcGIS 10*. ESRI Inc. [Online Help]. Online available: http://resources.arcgis.com/en/help/main/10.1/index.html#/An_overview_of_the_Interpolation_tools/009z00000069000000

Eumetcal 2013. Verification resource [online]. The European Virtual Organisation for Meteorological Training. Available from http://www.eumetcal.org/resources/ukmeteocal/verification/www/english/msg/ver_cont_var/uos3/uos3_ko1.htm [Accessed 30 October 2013]

Favalli M, Fornaciai A, Isola I, Tarquini S & Nannipieri L 2012 Multiview 3-D reconstruction in geosciences. *Computers & Geosciences* 44: 168-176.

Feetham LM, Aouf N, Bourdarias C & Voirin 2014. Single Camera absolute motion based on digital elevation mapping for next generation planetary lander. *Acta Astronautica* 98:169-188

Florinsky IV 1998. Accuracy of local topographic variables derived from digital elevation models. *International Journal of Geographical Information Science* 12(1): 47-62.

Fonstad MA, Dietrich JT, Courville BC, Jensen LJ & Carbonneau PE 2013. Topographic Structure from Motion: a new Photogrammetric Measurement. *Earth Surface Processes and Landforms* 38: 421-430

Forstner W 1986. A feature based correspondence algorithm for image matching. The International Archives of the Photogrammetry. *Remote Sensing* 26(3): 150-166

Frankl A, Stal C, Abraha A, Nyssen J, Rieke-Zapp D, De Wulf A & Poesen J 2015. Detailed recording of gully morphology in 3D through image-based modelling. *Catena* 127: 92-101.

Furukawa Y, Curless B, Seitz SM & Szeliski R 2010. *Towards internet-scale multi-view stereo. In Computer Vision and Pattern Recognition (CVPR).* Proceedings of IEEE Conference on Computer Vision and Pattern recognition held on 13-18 June. San Francisco: IEEE

Furukawa Y & Ponce J 2007. *Accurate, dense, and robust multi-view stereopsis*. Proceedings of the IEEE conference on Computer Vision held on 17-22 June 2007. Minneapolis: IEEE

Geomatica Orthoengine 2003. *User Guide 2003*. PCI Geomatica. [Online help] Online available: https://www.ucalgary.ca/appinst/doc/geomatica_v91/manuals/orthoeng.pdf

Ghosh SJ 2005. *Fundamentals of Computational Photogrammetry*. New Delhi: Concept Publishing Company

Golparvar-Fard M, Pena-More F & Savarese S 2011. *Monitoring changes of 3D building elements from unordered photo collections.* Proceedings of the Computer Vision workshop held 6-13 November 2011. Barcelona: IEEE international conference.

Gómez-Gutiérrez A, Schnabel S, Berenguer-Sempere F, Lavado-Contador F & Rubio-Delgado J 2014. Using 3D photo-reconstruction methods to estimate gully headcut erosion. *Catena* 120: 91–101.

Green S, Bevan A & Shapland M 2014. A comparative assessment of structure from motion methods for archaeological research. *Journal of Archaeological Science* 46: 173-181.

Hannah MJ 1989. A system for digital stereo image matching. *Photogrammetric Engineering & Remote Sensing* 55(12): 1765-1770.

Harwin S & Lucieer A 2012. Assessing the Accuracy of Georeferenced Point Clouds Produced via Multi-View Stereopsis from Unmanned Aerial Vehicle (UAV) Imagery. *Remote Sensing* 4: 1573-1599.

Heipke C 1996. Overview of Image Matching Techniques, in Proceedings of the OEEPE-Workshop on Application of Digital Photogrammetric Workstations held 4-6 March 1996. Lausanne

Hengl T & Reuter HI 2008. *Geomorphometry: Concepts, Software, Applications*. 1st ed. Oxford: Newnes

Hirschmuller H, Buder M & Ernst I 2012. *Memory efficient semi-global matching*. Proceedings of the 22nd ISPRS Congress held 25 August - 1 September 2012. Australia. Melbourne: International Society for Photogrammetry and Remote Sensing

Hodgson ME & Bresnahan P 2004. Accuracy of airborne LiDAR-derived elevation. *Photogrammetric Engineering & Remote Sensing* 70(3): 331-339.

Honkavaara E, Arbiol R, Markelin L, Martinez L, Cramer M, Bovet S, Chandelier L, Ilves R, Klonus S, Marshal P, Schläpfer D, Tabor M, Thom C & Veje N 2009. Digital Airborne Photogrammetry—A New Tool for Quantitative Remote Sensing?—A State-of-the-Art Review On Radiometric Aspects of Digital Photogrammetric Images. *Remote Sensing* 1: 577-605.

Hugenholtz CH , Ken Whitehead K, Brown OW, Barchyn TE, Moorman BJ, LeClair A, Riddell K & 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.

Hutchinson MF & Gallant JC 2000. Digital elevation models and representation of terrain shape. Terrain Analysis; Principles and Application. New York: John Wiley and Sons

James MR & Robertson S 2012. Straightforward reconstruction of 3D surfaces and topography with a camera: Accuracy and geoscience application. *Journal of Geological Research* [online] 117(3). Available from http://dx.doi.org/10.1029/2011JF002289 [Accessed 25 September 2014]

Jancosek M & Pajdla T 2011. *Multi-View Reconstruction Preserving Weakly-Supported Surfaces (CVPR)*. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition held 20-25 June 2011. Providence. Rhode Island: IEEE

Javernick L, Brasington J & Carusoa B 2014. Modelling the topography of shallow braided rivers using Structure-from-Motion photogrammetry. *Geomorphology* 213: 166–182.

Jensen JR 2007. *Remote Sensing of the Environment: An Earth Resource Perspective*. 2nd edition. Prentice Hall.

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

Küng O, Strecha C, Beyeler A, Zufferey JC, Floreano D, Fua P & Gervaix F 2011. The accuracy of automatic photogrammetric techniques on ultra-light UAV imagery. *Remote Sensing and Spatial Information Sciences* 38(1): 125-130.

Labatut P, Pons JP & Keriven R 2009. Robust and efficient surface reconstruction from range data. *Computer Graphics Forum* 28(8): 2275-2290.

Leberl F, Irschara A, Pock T, Meixner P, Gruber M, Scholz S & Wiechert A 2010.Point Clouds: Lidar versus 3D Vision. *Photogrammetric Engineering & Remote Sensing* 76(10): 1123-1134.

Lejot J, Delacount C, Piegay, Fourier T, Tremelo ML & Allemand P 2007. Very high spatial resolution imagery for channel bathymet ry and topography from an unmanned mapping controlled platform. *Earth Surfaces Processes and Landforms* 32: 1705-1725.

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

Lingua A, Marachino D & Nex F 2009. Performance Analysis of the SIFT Operator for automatic Feature Extraction and Matching in Photogrammetric Applications. *Sensors* 9(5): 3745-3766.

LPS Project Manager User Guide 2009. Oregon State University Repository. [Online help] Online available: http://classes.engr.oregonstate.edu/cce/fall2012/cce201-001/Photogrammetry/LPS_PM_classinstructions.pdf [Accessed 12 May 2015]

Macay Moreira JM, Nex F, Agugiaro G, Remondino F & Lim NJ 2013. From DSM to Building Models: A Quantitative Evaluation. *Remote Sensing and Spatial Information Sciences* 40: 213-219.

Mallet C & Bretar F 2009. Full-waveform topographic Lidar: State-of-the-art. *ISPRS Journal of Photogrammetry and Remote Sensing* 64(1): 1-16.

Matiukas V & Miniotas D 2011. Point cloud merging for complete 3D surface reconstruction. *Electronics and Electrical Engineering* 7(113): 73-76.

Meesuk V, Vojinovic Z, Mynett AE & Abdullah AF 2015. Urban flood modelling combining top-view LiDAR data with ground-view SfM observations. *Advances in Water Resources* 75: 105–117.

Micheletti N, Chandler JH & Lane SN 2015. Geomorphological Techniques. British Society for Geomorphology [online], Chap 2, Sec 2.2. Available from: http://www.geomorphology.org.uk/sites/default/files/geom_tech_chapters/2.2.2_sfm.pdf [Accessed 9 July 2015]

Mikolajczyk K & Scdmid C 2002. *An affine invariant interest point detector*. Proceeding of the 7th European conference on Computer Vision held on 28-31 May 2002. Denmark: ECCV

Mueller AR 2014. Lidar and Image Point Cloud Comparison. Master's thesis. Monterey: Naval Postgraduate School, Department of Information Science.

Muhlmann K, Maier D, Hesser J & Manner R 2002. Calculating Dense Disparity Maps from Color Stereo Images, an Efficient Implementation. *International Journal of Computer Vision* (47): 79-88

Myburgh G & Van Niekerk A 2013. Effect of Feature Dimensionality on Object-based Land Cover Classification: A Comparison of Three Classifiers. *South African Journal of Geomatics* 2(1): 13-27.

Nagl F, Kolzer K & Grimm P 2011. *Occlusion handling and image-based lighting using sliced images in 3D photo collections*. Proceedings of the 21st conference on Artificial Reality and Telexistence held 28-30 November 2011. Osaka: The virtual Reality Society of Japan.

Neithammer U, James MR, Rothmund S, Travelletti J & Joswig M 2012. UAV-based remote sensing of the super-sauze landslide: Evaluation and results. *Engineering Geology* 128: 2-11.

Newcombe RA & Davidson AJ 2010. Live Dense Reconstruction with a Single Moving Camera. IEEE 2010: 1498-1505.

Newcombe RA, Davidson AJ, Izadi S, Kohli P, Hilliges O, Shotton J, Molyneaux D, Hodges S, Kim D & Fitzgibbon A 2011. *Kinetic fusion: Real-time Dense surface Mapping and Tracking*. Proceedings of the IEEE International Symposium on Mixed and Augmented Reality held 26-29 October 2011. Basel, Switzerland. Congress Centre: IEEE

NGI 2014. Elevation Data [online]. National Geo-spatial Information. Available from http://www.ngi.gov.za/index.php/what-we-do/maps-and-geospatial-information/40-elevation-data [Accessed 30 October 2014]

NVidia 2013. NVidia Developer Zone [online]. CUDA Compute Capability Avaliable from https://developer.nvidia.com/cuda-gpus [Accessed 1 October 2013]

Oksanen J 2006. Digital Elevation Model Error in Terrain Analysis. Doctoral Dissertation. Helsinki: University of Helsinki, Department of Geography and Finnish Geodetic Institute

Okutomi M & Kanade T 1993. A multiple-baseline stereo. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 15(4):353–363

Photosynth 2005. Photosynth beta. Microsoft Corporation. [Online Help] Online available: http://photosynth.net/help.aspx.

Piantadosi S, Byar DP & Green SB 1988. The ecological fallacy. *American Journal of Epidemiology* 127(5), 893-904

Pomaska G 2009. *Utilization of Photosynth point clouds for 3D object reconstruction*. Proceedings of 22nd CIPA Symposium held 11-15 October 2009. Kyoto: CIPA.

Remondino F, Spera MG, Nocerino E, Menna F & Nex F 2014. State of the art in high density image matching. *The Photogrammetric Record* 29(146): 144–166.

Remondino F, Del Pizzo S, Kersten TP, Troisi S 2012. *Low-cost and open-source solutions for automated image orientation–A critical overview*. Proceedings of 4th International EuroMed Conference held October 29 – November 3 2012. Cyprus: Limassol

Rumsby BT, Brasing J, Langham JA, McLelland SJ, Middleton R & Rollinson G 2008. Monitoring and modelling particle and reach-scale morphological change in gravel-bed rivers: Applications and challenges. *Geomorphology* 39: 40-54.

Shapiro LG & Haralick RM 1987. Relational matching. *Applied Optics* 26 (10): 1845-1851

Sood B & Kaur K 2012. A monitoring approach for surface reconstruction from 3D point cloud. *International Journal of Computer technology & Applications* 3(1): 277-282.

Schmid C & Mohr R 1997. Local grayvalue invariants for image retrieval. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 19(5): 530-535.

Sate Information Technology Agency (SITA) 2007. Policy on Free and Open Source Software used for South African Government. Department of Public Service & Administration.

Seits SM, Curless B, Diebel J, Scharstein D & Szeliski R 2006. A Comparison and Evaluation of Multi-View Stereo Reconstruction Algorithms. *IEEE Conference on Computer Vision and Pattern Recognition* 1: 519-528.

Siebert S & Teizer J 2014. Mobile 3D mapping for surveying earthwork projects using an Unmanned Aerial Vehicle (UAV) system. *Automation in Construction* 41: 1–14.

Smith MW & 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: 1656-1671.

Snavely N, Seits SM & Szeliski 2006. Photo tourism: exploring photo collections in 3D. *ACM Transactions on Graphics* 25(3): 835-846.

Snavely N, Seitz SM & Szeliski R 2008. Modelling the world from internet photo collections. *International Journal of Computer Vision* 80: 198-210.

South African Civil Aviation Authority 2015. *Technical Guidance Material for RPAS – Part 101*. September. Pretoria: South African Civil Aviation Authority

Stal C, Bourgeois J, De Maeyer P, De Mulder G, De Wulf A, Goossens R, Hendrickx M, Nuttens T & Stichelbaut B 2012. *Quality Analysis of Structure from Motion in Archaeological Airborne Applications.* Proceedings of the 32th EARSeL Symposium: 'Advances in Geosciences' held 21-24 May 2012. Greece. Mykonos: European Association of Remote Sensing Laboratories

Stal C, De Wulf A, De Maeyer P & Goossens & Nuttens T 2012. *Evaluation of the accuracy of 3D data acquisition Techniques for the documentation of cultural Heritage*. Proceedings of the 3rd International EARSeL Workshop on the Advances in Remote Sensing for Archaeology and Cultural Herritage Management held on 19-22 September 2012. Ghent, Belgium. Ghent: EARSeL

Strecha, C, Bronstein A, Bronstein M & Fua P 2011, LDAHash: Improved matching with smaller descriptors. *IEEE Pattern Analalysis and Machine Intelligence* 34: 66–78.

Tavani S, Granado P, Corredetti A, Girundo M, Lannace A, Arbues P, Munoz JA & Mazzoli S 2014. Building a virtual outcrop, extracting geological information from it, and sharing the results in Google Earth buyer open pilot and photo scan: an example from the Khaviz Anticline (Iran). *Computers & Geosciences* 63: 44-53.

Thomson WT 1965. *Vibration theory and applications*. Cheltenham: Prentice-Hall

Tonkin TN, Midgley NG, Grahamb DJ & Labadz JC 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.

Torres JC, Arroya G, Romo C & De Haro J 2012. *3D digitisation using Structure from Motion*. Proceedings of Spanish Computer Graphics Conference held 12-14 September 2012. Jaén, Spain: CEIG.

Tsai RY 1983. Multiframe image point matching and 3-d surface reconstruction. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 5(2):159–174.

Tsay JR & Lee MS 2012. Sift for Dense Point Cloud Matching and Aero Triangulation. *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences* 39: 69-74.

Turner D, Lucieer A & Watson C 2012. An automated technique for generating georeferenced mosaics from ultra-high resolution unmanned aerial vehicles (UAV) imagery based on structure from motion (SfM) point clouds. *Remote Sensing* 4: 1392-1410.

Van der Merwe D & Price KP 2015. Harmful Algal Bloom Characterization at Ultra-High Spatial and Temporal Resolution Using Small Unmanned Aircraft Systems. *Toxins* 7(4): 1065-1078.

Van Niekerk A 2012. Developing a very high resolution DEM of South Africa. *PositionIT* November/Dececember: 55-60.

Vasuki Y, Holden EJ, Kovesi P & Micklethwaite S 2014. *A geological structure mapping tool using photogrammetric data.* 23rd International Geophysical Conference and Exhibition held 11-14 August 2013. Australia:Melbourne

Wackrow R & Chandler JH 2008. A convergent image configuration for DEM extraction that minimises the systematic effects caused by an inaccurate lens model. *The Photogrammetric Record* 23(121): 6-18

Wang CK, Tseng YH & Chu HJ 2014. Airborne Dual-Wavelength LiDAR Data for Classifying Land Cover. *Remote sensing* 6: 700-715.

Wang Y 1998. Principles and Applications of Structural Image Matching. *ISPRS Journal of Photogrammetry and Remote Sensing* 53: 154-165.

Warren M, Mckinnon D, He H, Glover A, Chiel M & Upcroft B 2014. *Large Scale Monocular Vision-only Mapping from Fixed wing sUAS*. In the Field of Robotics: Results of the 8th international Conference 2014: 495-509.

Westoby MJ, Brasington J, Glasser NF, Hambrey MJ & Reynolds 2012. 'Structure-from-motion' photogrammetry: A low-cost, effective tool for geoscience applications. *Geomorphology* 179: 300-314.

Wu C 2011. VisualSfM 5.22. VisualSFM: A Visual Structure from Motion System. Washington University. [Online] available from http://homes.cs.washington.edu/~ccwu/vsfm/ [Accessed 12 August 2013]

Wu C 2013. Towards Linear-time Incremental Structure from Motion. *International Conference on 3D Vision* 2013: 127-134.

Zhang Y 2002. Problems in the fusion of commercial high-resolution satellite as well as Landsat 7 images and initial solutions. *ISPRS GeoSpatial Theory, Processing and Applications* 34(4): 587-592

Zhou Q & Liu X 2004. Error analysis on based slope and aspect algorithms. *Photogrammetric Engineering and Remote Sensing* 70(8): 957 - 962.

Zhou W 2013. An Object-Based Approach for Urban Land Cover Classification: Integrating LiDAR Height and Intensity Data. *IEEE Geoscience and Remote Sensing Letters* 10(4): 928 – 931.

PERSONAL COMMUNICATIONS

Luck-Vogel M 2015. Senior Researcher Coastal Remote Sensing. Stellenbosch. Email on 5 October about data acquisition technique considerations- 2014 evaluation.