# Geodätisch-geophysikalische Arbeiten in der Schweiz

(Fortsetzung der Publikationsreihe
«Astronomisch-geodätische Arbeiten in der Schweiz»)

herausgegeben von der

Schweizerischen Geodätischen Kommission
(Organ der Akademie der Naturwissenschaften Schweiz)

**Neunundneunzigster Band**
**Volume 99**

# Monitoring Rock Glaciers by Combining Photogrammetric and GNSS-Based Methods

Fabian Neyer

2017

Adresse der Schweizerischen Geodätischen Kommission:

Institut für Geodäsie und Photogrammetrie
Eidg. Technische Hochschule Zürich
ETH Zürich
8093 Zürich
Switzerland

Internet: http://www.sgc.ethz.ch

ISBN 978-3-908440-45-1

Redaktion des 99. Bandes:
Dr. F. Neyer, J. Müller-Gantenbein, Prof. A. Geiger
Druck: Print-Atelier ADAG, Zürich

# VORWORT

Das Monitoring von Naturgefahren gewinnt an Bedeutung, da die Risikobereitschaft sinkt und das Schadenspotential mit zunehmender Verbauung steigt. Mit dem gegenwärtigen Trend der Klimaerwärmung geht auch ein steigendes Gefahrenpotential von potentiellen Rutschgebieten aus. Insbesondere sind Permafrost Areale vom Auftauen bedroht und können dadurch ihre Stabilität verlieren.
In diesem Kontext ist auch die vorliegende Publikation von Herrn Neyer zu sehen. Er hat sich zum Ziel gesetzt, ein quasi-automatisches Werkzeug zur genauen Überwachung von Rutschungen bereitzustellen. Mit der entsprechenden Beobachtungsmethode sollten auch kleine Terrain-Bewegungen festgestellt werden können, dies ohne viele manuelle Interventionen zu verlangen. Die hier gewählten Ansätze basieren auf der Kombination bildverarbeitender Methoden mit GNSS Messungen.
Die Untersuchungen und Entwicklungen von Herrn Neyer sind in das X-Sense Projekt eingebettet, das unter anderem der Entwicklung von Monitoringsystemen zur Überwachung von Blockgletschern gewidmet ist. Die Messsysteme sollen eine hohe zeitliche und räumliche Auflösung liefern und möglichst autonom operieren. Die mehrjährigen Beobachtungssequenzen sollen schliesslich zu einem besseren Prozessverständnis der Rutschvorgänge beitragen. Dabei bilden die Automatisierung und Optimierung der Datenverarbeitung eine besondere Herausforderung insbesondere im Bereich der Bildverarbeitung. Das Projekt 'x-sense' wurde im Rahmen des SNF Forschungsprogramms 'Nano-Tera' gefördert und betraf die 'Technische Informatik und Kommunikation' der ETH Zürich, die 'Geodäsie und Photogrammmetrie', ETH Zürich sowie die 'Physische Geographie' der Uni Zürich. Als Partner beteiligt waren auch Gamma Remote Sensing und das BAFU.

Diese Untersuchung reiht sich in die Arbeiten zum Geomonitoring des Institutes für Geodäsie und Photogrammmetrie der ETH Zürich und der Schweizerischen Geodätischen Kommission (SGK) ein. Wir danken dem Verfasser, Herrn Neyer, für den wertvollen Beitrag zum geodätischen Geomonitoring. Dem SNF mit seinem 'Nano-Tera'-Forschungsprogramm gebührt Dank für die Teilfinanzierung und Unterstützung, genauso wie dem BAFU für die finanzielle und technische Hilfeleistung.
Der SCNAT danken wir für die Übernahme der Druckkosten.

**Prof. Dr. M. Rothacher**
Institut für Geodäsie und Photogrammetrie
ETH Zürich

**Prof. Dr. Alain Geiger**
ETH Zürich
Präsident der SGK

**PREFACE**

Le monitorage des dangers naturels gagne en utilité et en intérêt. En effet, l'acceptation des risques devient en général plus faible, et les dégâts potentiels augmentent avec l'extension du tissu construit. Avec le réchauffement climatique actuel, les accidents potentiels liés aux zones soumises à des glissements de terrain s'accroissent. En particulier, le dégèle des zones en permafrosts peuvent provoquer une perte significative de la stabilité des sols sous-jacents.

C'est dans ce contexte que se présente la publication de Monsieur Neyer. Il s'était fixé comme objectif de créer un outil de surveillance de glissement de terrain quasi-automatique. Grâce à la technologie mise en place, qui s'appuie sur des techniques de photogrammétrie et de GNSS, de petits mouvements de terrains devraient pouvoir être détectés. Et ceci, en limitant les interventions manuelles au strict minimum.

Les études et les développements de Monsieur Neyer font partie du projet X-Sense, et en particulier du développement des systèmes de monitoring pour la surveillance de glaciers rocheux. Le système doit permettre de produire des données à hautes résolutions spatiale et temporelle en opérant de manière la plus automne possible. Les séquences d'observations qui en résultent peuvent ainsi contribuer à une meilleure compréhension des processus physique en jeu lors de glissements de terrain. A ce propos, l'automatisation de la chaîne des traitements des données, et en particulier celle des traitements d'images, contient des défis techniques majeurs. Le projet X-Sense a été soutenu par le programme de recherche Nano-Tera du FNS et fût réalisé par les Instituts d'informatique technique et communication de l'EPF de Zürich, de géodésie et de photogrammétrie de l'ETH de Zürich ainsi que de celui de géographie physique de l'Université de Zürich. De plus, l'entreprise Gamma Remote Sensing ainsi que l'OFEV ont officié comme partenaire.

Le présent travail représente une pièce maitresse dans la série des recherches sur le géomonitorage de haute précision de l'IGP et de la commission géodésique suisse (CGS). Nous remercions Monsieur Neyer pour cette contribution de grande valeur à la géodésie. Nos remerciements vont aussi au FNS qui, à travers son Programme de recherche 'Nano-Tera' a fortement co-financé le projet et à l'OFEV pour son support actif et le financement partiel de ces recherches.

La Commission Géodésique Suisse (CGS) est reconnaissante envers l'Académie Suisse des Sciences Naturelles (SCNAT) pour avoir pris à sa charge les coûts d'impression du présent manuscrit.

**Prof. Dr. M. Rothacher**
Institut de Géodésie et Photogrammétrie
ETH Zürich

**Prof. Dr. Alain Geiger**
ETH Zürich
Président de la CGS

## FOREWORD

The monitoring of natural hazard increasingly gains of attention since awareness of risk and the value of build environment is steadily growing. The present trend of climatic warm-up converts possible sliding areas into potentially hazardous areas. More specifically permafrost areas are prone to melting and thus to loss of stability.

The present publication of Mr. Fabian Neyer has to be seen in this context. He aimed at an optical tool to automatically monitor and survey terrain slides at high precision. The devised observation method should also detect small movements without requiring manual interventions. The principles of the method are based on the combination of imaging and GNSS.

Mr. Neyer's investigations and developments are embedded in the 'x-sense'-project that is dedicated, among others, to the development of monitoring systems of rock glaciers. The systems shall provide a high spatio-temporal resolution as well as a high level of autonomy. The observation over a time span of several years shall lead to an enhanced understanding of ongoing sliding processes. To this end the automation and optimization of the data treatment especially in image processing was a crucial element. The project 'x-sense' has been supported in the frame of the SNF-Program 'Nano-Tera' and has involved the 'Technical Informatics and Communication' at ETH Zurich, the 'Institute of Geodesy and Photogrammetry', ETH Zurich, and the 'Physical Geography' of Uni Zurich. Partners were Gamma Remote Sensing and FOEN.

This investigation represents a further master piece in the series of geomonitoring research of the Institute of Geodesy and Photogrammetry and the Swiss Geodetic Commission (SGC).

Thanks go to the author, Fabian Neyer, for his valuable contribution to geodetic geomonitoring.

We thank the SNF who, through its program 'Nano-Tera', strongly supported this work as well FOEN, whose financial and technical support is greatly appreciated. Thanks are given to the Swiss Academy of Sciences for covering the printing costs of this volume.

**Prof. Dr. M. Rothacher**
Institute of Geodesy and Photogrammetry
ETH Zurich

**Prof. Dr. A. Geiger**
ETH Zurich
President of SGC

# Abstract

Rock glaciers are creeping landforms of perennially frozen ground and belong to the permafrost creeping phenomena. They are mainly composed of rock debris that accumulate in areas of high natural erosion. Ice particles between the rocks cause the moving accumulation in steep terrain to dynamically flow downslope. In the Alpine region, these morphological landforms mainly occur at north-facing mountain slopes in high altitudes above the forest boundary and are known for their sensitivity to climate change.

For several decades, rock glaciers have been monitored for scientific aims, while advances in surveying technologies increased the interest in such studies since the 1990s. Modern technologies in remote sensing (e.g., airborne imagery or satellite-based measurement techniques) are often combined with measurements from field campaigns, i.e., measurements taken directly on a rock glacier (e.g., GNSS, laser-scanning, ground temperature measurements, etc). The high-level goal is to enhance the process understanding, specially with respect to the changing climate: various studies indicate an extended risk of slope failures in steep frozen bedrock due to the global temperature increase. Early recognition of increased activities help to inform local authorities in the endangered areas about the potential hazard before such an event.

The present work is part of the *X-Sense* project (Nano-Tera.ch), with an interdisciplinary team of scientists that build and operate new low-cost devices for data acquisition, develop new data processing pipelines and algorithms for evaluation, and also gain new insight of natural processes in these regions. Autonomous measurement systems, developed within other work packages in the *X-Sense* project, observe different permafrost creep areas with high resolution in space and time. Combined with multi-year observations, the derived surface motions are used to obtain an improved process understanding.

This work focuses on the photogrammetric image processing in order to retrieve precise surface displacement estimates. More precisely, image sequences, acquired with two permanently installed commercial digital single-reflex cameras, are used to measure topographic changes in the observed permafrost area. By the combination with high resolution GNSS positioning results, the goal is to obtain precise time series of moving rock boulders at different positions within the field of view. Challenges arising from the combination of different data sets, the development of an automatic processing pipeline, and an improvement of the processing strategy in general, are the main tasks of this thesis.

The study site is the bordering area above the Grabengufer rock glacier (Mattervalley VS, Switzerland), known as the Grabengufer rock slide. Local topographic conditions allowed only a partially good installation geometry for the photogrammetric reconstruction. With respect to a 3D reconstruction without the use of GNSS coordinates, an accuracy increase of about one order of magnitude could be achieved in case these high-precision solutions were integrated. More specifically, respective standard deviations for the East, North, and Height components of 6, 5, and 2 cm were achieved. The stated accuracy, maintained throughout the measurement period of nearly four years (summer months), was obtained in an area of approximately 80 m × 80 m, with a mean distance of 80 m from the two cameras.

Position time series of moving rock boulders were filtered using the principles of collocation. Analyzing the correlation characteristics of the stochastic signal, an optimal correlation length was computed and used to extract relevant signals from the noise contaminated time series. Velocity was directly estimated as a derived quantity in the collocation process. Furthermore, the techniques of the adaptive collocation approach is presented. This iterative method uses the principles of a dynamically adjusting anisotropic covariance metric. In an example of 2-dimensional velocity fields it is shown that regional compression and extension areas can be extracted.

Results indicate that the observed permafrost area has experienced a mean annual acceleration of about 0.1 m/year between the years 2013 and 2015. During the late summer months of 2015, a prominent temporal acceleration was observed. The mean displacement rate was found to be 0.67 m/year, whereas the 3-dimensional displacement is dominated by a translation following the gliding surface. An area in the front of the observed field of view was found to have higher displacement rates, specially during the late summer months, thus it detaches from the otherwise relatively homogeneous flow field.

The methods and principles presented in this work show the potential of monitoring permafrost surface displacements using permanently installed optical cameras in combination with positioning results from permanently mounted GNSS stations. These principles can easily be transfered to other monitoring applications and thus contribute to a better understanding of such processes.

# Zusammenfassung

Blockgletscher sind kriechende Schuttmassen in Gebieten mit mehrjährig gefrorenem Untergrund und gehören somit zu den Permafrost-Kriechphänomenen. Sie bestehen im wesentlichen aus Gesteinsmaterial, welches durch erhöhte natürliche Erosionsprozesse angehäuft wurde. Eispartikel zwischen den Gesteinsbrocken bewirken, dass sich diese Masse im steilen Gelände dynamisch talabwärts bewegt. Oft sind Blockgletscher unterhalb von Endmoränen bei Eisgletschern anzutreffen. Im Alpenraum sind sie hauptsächlich an nach Norden ausgerichteten Berghängen in grossen Höhen oberhalb der Waldgrenze zu finden. Sie sind als empfindliche Indikatoren für den Klimawandel bekannt.

Seit einigen Jahrzehnten werden Blockgletscher wissenschaftlich untersucht, wobei Fortschritte in der Messtechnik seit den 1990er Jahren zu einer starken Zunahme solcher Studien geführt haben. Moderne Techniken der Fernerkundung (z.B. Luftaufnahmen oder satellitengestützte Messtechniken) werden oft mit Messungen von Feldkampagnen, d.h. Messungen vor Ort auf den Blockgletschern (z.B. GNSS, Laserscans, Temperaturmessungen etc.) ergänzt. Das übergeordnete Ziel ist ein besseres Prozessverständnis, speziell im Zusammenhang mit dem sich ändernden Klima: Diverse Studien deuten darauf hin, dass Destabilisierungsprozesse in steilen Berghängen mit Dauerfrost, ausgelöst durch die globale Temperaturzunahme, zu erhöhten Risiken von Hangrutschungen führen. Die frühzeitige Erkennung solcher Abläufe ermöglicht es, die Behörden der betroffenen Gebiete rechtzeitig über die potentielle Gefahr zu informieren.

Die vorliegende Arbeit ist Teil des Projektes *X-Sense* (Nano-Tera.ch), in welchem ein interdisziplinäres wissenschaftliches Team neue sogenannte 'low-cost' Sensoren baut und betreibt, neue Algorithmen zur Datenverarbeitung entwickelt und daraus neue Erkenntnisse über die ablaufenden Prozesse in diesen Regionen erhält. Autonome Messsysteme, entwickelt in anderen Arbeiten innerhalb des *X-Sense* Projektes, werden zur Messung diverser Permafrost-Kriechprozesse mit hoher zeitlicher und räumlicher Auflösung eingesetzt. Mittels mehrjähriger Beobachtungen werden die damit errechneten Oberflächenverschiebungen für ein verbessertes Prozessverständnis genutzt.

Schwerpunkt dieser Arbeit ist die photogrammetrische Bildverarbeitung in Bezug auf die präzise Messung von Oberflächenverschiebungen. Konkret sollen Bildsequenzen, aufgenommen mit zwei permanent installierten, kommerziellen Spiegelreflexkameras, für die Vermessung von Oberflächenveränderungen im beobachteten Permafrostgebiet genutzt werden. Durch Kombination mit zeitlich hochaufgelösten GNSS-Positionsmessungen sollen präzise Zeitserien von sich bewegenden Steinblöcken an unterschiedlichen Positionen innerhalb des beobachteten Bereiches berechnet werden. Zu den Hauptaufgaben dieser Arbeit gehören die Bewältigung von Schwierigkeiten bei der Kombination verschiedener Datensätze, die

Automatisierung der Informationsverarbeitung sowie eine optimierte Verarbeitungsstrategie im Allgemeinen.

Standort der Studie ist das angrenzende Gebiet oberhalb des Grabengufer Blockgletschers (Mattertal, VS, Schweiz), auch bekannt als die Felsrutschung am Grabengufer. Lokale Gegebenheiten führten zu einer Installationsanordnung, welche für die photogrammetrische Auswertung als nur teilweise gut bezeichnet werden kann. Im Vergleich zur 3D Rekonstruktion ohne Nutzung präziser GNSS Positionen konnte durch deren Integration eine Genauigkeitssteigerung von etwa einer Grössenordnung erreicht werden. Konkret wurden für die Komponenten Ost, Nord und Höhe Standardabweichungen von 6, 5 und 2 cm erreicht. Diese Messgenauigkeit, geltend für ein ca. 80 m × 80 m grosses Gebiet in einer mittleren Entfernung von 80 m von den beiden Kamerastandorten, konnte über die gesamte Messperiode von knapp vier Jahren (Sommermonate) beibehalten werden.

Zur Filterung der gewonnenen Positionszeitreihen von Gesteinsblöcken wurden die Prinzipien der Kollokation angewandt. Mittels Analyse des stochastischen Signals auf Korrelation wurden optimale Korrelationslängen bestimmt. Diese wurden genutzt, um relevante Signale aus den mit Rauschen behafteten Zeitreihen zu erhalten. Die Geschwindigkeit wurde direkt als abgeleitete Grösse im Kollokationsprozess mitbestimmt. Des Weiteren wird die Technik der adaptiven Kollokation vorgestellt. Die iterative Methode nutzt das Prinzip einer sich dynamisch anpassenden, anisotropen Korrelationsmetrik. Im Beispiel von 2-dimensionalen Verschiebungsfeldern wird gezeigt, wie sich damit auch regionale Kompressions- und Extensionsgebiete bestimmen lassen.

Die Resultate zeigen, dass sich das beobachtete Permafrostgebiet zwischen 2013 und 2015 im Mittel um jährlich ca. 0.1 m/Jahr beschleunigt hat. Im Spätsommer 2015 wurde zudem eine ausgeprägte temporäre Zunahme der Verschiebungsgeschwindigkeit festgestellt. Die durchschnittliche Verschiebungsrate beträgt 0.67 m/Jahr, wobei die 3-dimensionale Verschiebung als eine dem Rutschhorizont folgende Translation beschrieben werden kann. Ein Gebiet im Frontbereich des beobachteten Ausschnitts zeigt eine erhöhte Verschiebungsrate, speziell in den späten Sommermonaten, was zu einer Ablösung von dem sonst relativ homogenen Verschiebungsfeld führt.

Die Methoden und Prinzipien, welche in dieser Arbeit präsentiert werden, zeigen das Potential der Überwachung von Oberflächenverschiebungen in Permafrostgebieten mittels permanent installierten optischen Kameras in Kombination mit Positionslösungen von permanent aufgebauten GNSS Stationen. Diese Prinzipien lassen sich auch leicht auf andere Überwachungsaufgaben anwenden und tragen so zum besseren Verständnis solcher Phänomene bei.

# Contents

# Abbreviations

| | |
|---|---|
| ALSC | Adaptive Least-Squares Collocation |
| CEST | Central European Summer Time |
| DSLR | Digital Single-Lens Reflex (camera) |
| DEM | Digital Elevation Model |
| DInSAR | Differential Interferometric Synthetic Aperture Radar |
| FOEN | Federal Office of the Environment |
| FOV | Field of View |
| GNSS | Global Navigation Satellite System |
| GPS | Global Positioning System |
| InSAR | Interferometric Synthetic Aperture Radar |
| LSC | Least-Squares Collocation |
| LSM | Least-Squares Matching |
| SNSF | Swiss National Science Foundation |

# 1 Introduction

## 1.1 Background and Motivation

Long term temperature variations, precipitation, wind, and other erosion processes are natural phenomena leading to a continuously changing surface topography throughout Earth's history (*Press and Siever*, 2001). With a rapid increase of temperature due to climate change, glaciers retreat at high speeds and permanently frozen ground (permafrost[1]) starts to warm up. Permafrost regions are present not only at high latitudes but also in high altitudes, e.g., in the Alps (*Häberli*, 1985). Although not the only controlling force, permafrost is known to affect slope instabilities (e.g. *Noetzli et al.*, 2006), specially in steep bedrock (*Gruber and Haeberli*, 2007). With a permanently increasing population, more infrastructure will become affected or endangered by natural hazards with small to catastrophic scales. Besides the scientific interest, it is thus important to study the influence between the changing climate and slope instabilities in alpine regions. A good understanding of such interactions help to build, e.g., suitable protection monuments or to build reliable early warning systems in cases of acute danger.

Rock glaciers are creeping landforms in permafrost areas, typically associated with perennially frozen rocks. It is currently assumed that the number of active landforms is in the order of 2'000 for the Swiss Alps alone (*Delaloye et al.*, 2010). This also includes rock glaciers that are - due to a high correlation of their motion rates with mean annual temperatures - often described as key landforms for the study of permafrost creep (e.g., *Barsch*, 1996). Temperature, however, is not the only quantity influencing the displacements: Local erosion rates (sedimentary loading), slope steepness, bedrock topography, and hydrological conditions are also key parameters that control the motion behavior. Given a relatively constant average precipitation rate for a given region, changing creep characteristics of these landforms are most likely linked to the global temperature change. Such a relation is also shown by acceleration periods correlating with increased temperatures observed at different rock glaciers throughout the Alps (*Delaloye et al.*, 2008a). As correlation with temperature seems to be confirmed, a detailed understanding of processes leading to the destabilization of rock glaciers and permafrost creep in general is missing.

The very principle of process understanding is based on various quantitative and sometimes also qualitative observations. Long-term rock glacier kinematics are typically estimated using scanned aerial photographs for epochs before the digital era (e.g., *Kaufmann*, 1998) and historical knowledge may be obtained by local inhabitants (e.g., *Delaloye et al.*, 2013). Since the early 1990s, rock glaciers are more intensively studied using modern surveying

[1]Defined as ground frozen for at least two consecutive years

techniques: Individual rock boulders are monitored using the principles of the Global Positioning System (GPS) or more generally the Global Navigation Satellite System (GNSS[2]), either by repeated measurements (*Lambiel and Delaloye*, 2004) or with permanent low-cost installations over long periods (*Limpach and Grimm*, 2009). Measurement techniques such as InSAR (synthetic aperture radar interferometry) (*Delaloye et al.*, 2007; *Lambiel et al.*, 2008) or aerial photographs (e.g., *Kääb*, 2002) are frequently used to obtain displacement maps over extended areas.

It is well known that every measurement system has its strengths and weaknesses, either due to the measurement technique itself or due to the geometrical configuration during operation. Modern research aims to combine complementary surveying techniques typically leading to improved accuracies and resolution. When applied with current state-of-the-art algorithms and processing strategies, GNSS-based positioning with low-cost receivers (as used in this study) can achieve accuracies in the centimeter to millimeter range (*Limpach and Grimm*, 2009). The typical accuracy in photogrammetric reconstruction, on the other hand, is worse by at least one order of magnitude, depending on the geometrical setting and camera configuration (e.g., focal length, object distance etc.) (e.g., *Kaufmann and Seier*, 2016). The strength of GNSS is to estimate a precise position of the antenna center in space and time. To a large extend this method does not dependent on the weather condition, day, or night. A simultaneous estimation of various positions seen by multiple cameras is the strength of the photogrammetric method. Limited by optical visibility, however, this method can only be applied during daytime and with targets (i.e., rocks) being visible, i.e., not covered by snow or hidden behind clouds and other obstacles in the line-of-sight.

The aim of this study is to apply and develop state-off-the-art processing algorithms to reliably reconstruct the motion of various rock boulders seen in image sequences. Combining the photogrammetric processing methodology with high accuracy GNSS coordinates increases the valuable surface motion estimates during the summer months. An areal coverage of precise surface motion contributes to a better understanding of permafrost creep with respect to the changing environmental conditions.

The Fourth Assessment Report of the International Panel on Climate Change (IPCC) has drawn worldwide attention to the future impact of global warming and its regional consequences (*Alcamo et al.*, 2007; *Christensen et al.*, 2007). The highly sensitive and often non-linear response of cryosphere phenomena to climate change, as well as important feedback mechanisms, make the understanding of cryosphere systems a high priority for climate and climate impact research. It is also conceivable that the monitoring and modeling capabilities, e.g., the proposed data combination strategy developed in this thesis, can be adapted and transferred to other domains and hazard scenarios, e.g., flooding, rockfall, or the protection of purposely built infrastructure. Switzerland and the alpine arc in general exhibit a high density of various scientific measurements, making it an ideal testbed for methods and technologies to be transferred to other mountain ranges worldwide.

[2] GNSS in general can include GPS (American system), GLONASS (Russian system), BeiDou (Chinese system), and GALILEO (European system) satellite data.

## 1.2 The *X-Sense* Project

This study is part of the *X-Sense* project that tackles the development of equipment and methods for measuring rock glacier behavior across different scales in space and time. The project is financed by Nano-Tera.ch[3], a program of SNSF (Swiss National Science Foundation). The *X-Sense* consortium is led by Prof. L. Thiele, TIK, ETHZ. Using a set of novel hard- and software tools, combined with real-time measurements at the field site and near-real-time data processing, the foundation of an early warning system with todays state-of-the-art products, is obtained. The measurement principle and system describtion of devices used in this project are given in *Beutel et al.* (2011). *Limpach et al.* (2011) describes the potential of the low-cost GPS devices used within *X-Sense*. In *Wirz et al.* (2016), an overview of the various sites and displacement results are presented.

Various projects assessing risks, e.g., from large volume landslides, show the importance of interdisciplinary studies of such potential hazards. Concentrating the various scientific expertises from hardware development, electronic engineering, geodesy, remote sensing, hydrology, geomorphology, etc., new insights in complex mechanisms of such landforms may be obtained. The tools and methods developed in this project can be used, e.g., for studies on permafrost creep dynamics over time scales from days to years, the study of correlation between permafrost creep (e.g., rock glaciers) and environmental changes such as changing annual temperatures, precipitation, and more. The application orientated project is a collaboration among the following parties:

- The Computer Engineering Group, TIK, ETH Zürich. Principle investigator. The group develops, tests, deploys, and operates new prototype low-cost sensors that are build for operation in the alpine environment. Data gathered with these instruments are the RAW data sources of the studies conducted within the *X-Sense* project.
- The group of Mathematical and Physical Geodesy, IGP, ETH Zürich. Here the task is to optimally process the GPS/GNSS and camera data received from the sensors, i.e., GPS raw data is processed for accurate positioning results and camera images are used to retrieve accurate surface motion over an extended area (presented in this thesis).
- Gamma Remote Sensing. A partner company that develops and applies InSAR processing routines to reveal surface height changes over large areas.
- The Glaciology and Geomorphodynamics Group, Physical Geography, University of Zürich. Their task is to interpret the displacement estimates obtained so far. Also local surface temperatures and other meteorological data is evaluated for this task.
- The Federal Office for the Environment (FOEN). FOEN is a financial supporter, a geological advisor, and the primary customer.

Using low-cost sensors as primary measurement devices, the *X-Sense* project aims to establish a permanent monitoring setup in high alpine altitudes to 'sense' various physical and environmental processes. The content of this thesis with respect to the project is also illustrated in Fig. 1.1.

[3] http://www.nano-tera.ch/projects/227.php

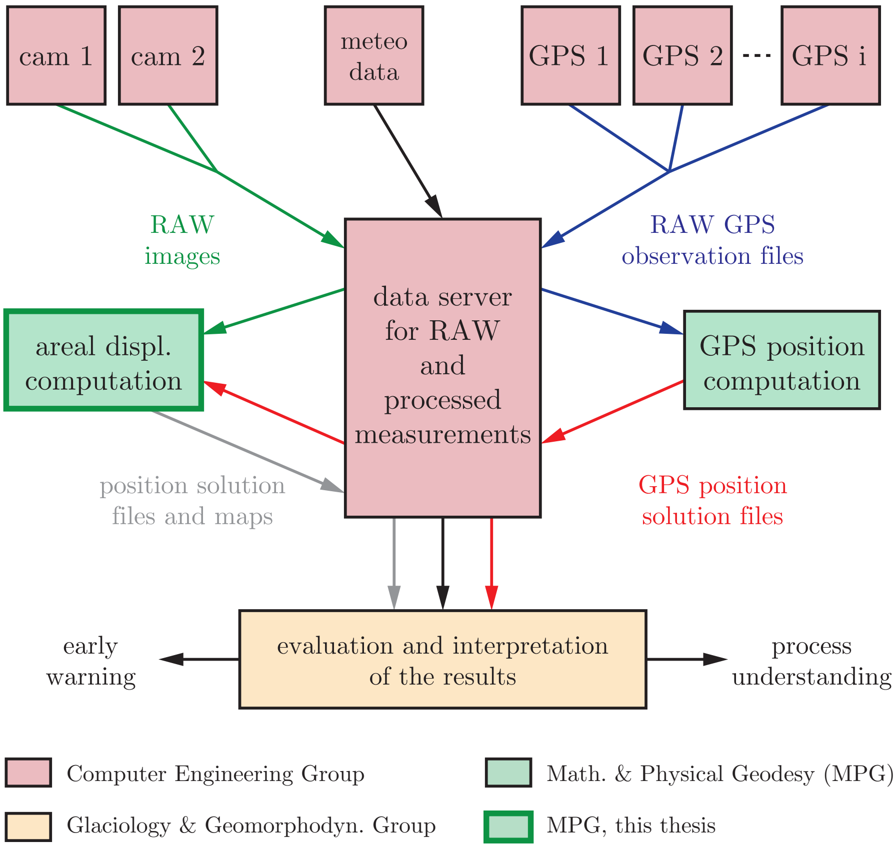


**Figure 1.1:** *X-Sense* contribution overview. The various RAW measurements, i.e., RAW images, a large number of RAW GPS signals (GPS 1, 2, ..., i), meteorological measurements, and more, is provided by the Computer Engineering Group. Their work includes the development, deployment, and operation of the various measurement devices. Also they are responsible for data transfer, data hosting, data accessibility and storage. Validation of the measurements and the processing of positioning solution time series is accomplished by the group of Mathematical and Physical Geodesy. The combination of the various results and measurements for characterizing the permafrost creep behavior is done in the group of Glaciology and Geomorphodynamics. The grey arrows indicate a solution transfer that has yet to be implemented.

## 1.3 State-of-the-Art and Challenges

To study the kinematics and dynamics of permafrost creep, continuous high accuracy measurements are of importance (e.g., *Wirz et al.*, 2016). Typical tools used to obtain such information are InSAR measurements (*Kilburn and Petley*, 2003; *Delaloye et al.*, 2007, 2008b; *Lambiel et al.*, 2008; *Strozzi et al.*, 2008), DInSAR (differential InSAR) (*Strozzi et al.*, 2004), repeated airborne imagery (e.g., *Kääb*, 2002), terrestrial or air-born laser scanners (*Travelletti et al.*, 2008; *Bell et al.*, 2012), or GPS/GNSS measurements of selected rock boulders (*Limpach and Grimm*, 2009).

InSAR and DInSAR are frequently applied to detect creeping landforms, measuring line-of-sight displacements in the centimeter to millimeter range with a spatial resolution of a few meters. A great advantage of this remote sensing technology is its large areal coverage also for regions with limited accessibility. Spatial coverage, however, is reduced to some extend by layover and shadowing effects of the rugged topography in permafrost areas. Resolution in time is given by the repeatability of the satellite flyovers and range from 2.5 days (TerraSAR-X) to 35 days (ERS-2) but can be increased by combining two images from any SAR mode. Ground based radar interferometry is another fast developing technique that complements the satellite based techniques (e.g. *Strozzi et al.*, 2015). As in the case of optical imagery, ground- and satellite-based InSAR is successfully applied in the absence of snow. Snow coverage results in decorrelation such that displacements cannot be estimated (*Lundgren*, 2009).

Terrain or construction deformation monitoring by GNSS is a technology that has been developed to highly performing systems, e.g., Leica (*Geosystems*, 2009). As an on-line reference for surveying purposes, many countries have set up a permanent GNSS network (e.g., *Wild et al.*, 2006). Specialized companies have established solutions for small scale monitoring investigations that are provided by GNSS measurements or by terrestrial geodetic equipment, e.g. automated theodolites and leveling equipment (*Blaikie et al.*, 2004; *Schätti and Rub*, 2008; *Solexperts*, 2009). With a typical accuracy of 0.5 cm for daily solutions and 1.5 cm for kinematic solutions (*Limpach and Grimm*, 2009) this method has proven to be one of the most accurate and robust monitoring tools available today. Being operational also in bad weather conditions, various GNSS-based applications have been developed for different tasks. For example early warning of volcanic activities (*Lundgren*, 2009), the detection of rapid and small movements (*Guillaume and Geiger*, 2007), or the determination of slow deformation patterns (*Hollenstein et al.*, 2006, 2008a,b)), to name just a few.

Terrestrial photogrammetry (e.g., *De Matías et al.*, 2009; *Kaufmann*, 2012) and airborne imagery (e.g., *Kääb et al.*, 1997) in general were successfully applied for studies of rock glacier flow fields over various time intervals. The method of normalized cross-correlation for matching features between different epochs in an image sequence is often suggested as a very fast and easy-to-use procedure for flow estimation with accuracies in the order of one pixel (*Kääb and Vollmer*, 2000). Other application orientated projects with aerial photographs (e.g., *Kaufmann and Ladstädter*, 2003) rely on the principle of least-squares image matching that is based on a more flexible mathematical model with intrinsic subpixel accuracy.

*Akca* (2013) showed the potential of photogrammetric point estimation for a small landslide triggered in a test environment. Using four cameras in geometrically optimized positions, signalized target points were tracked frame by frame and their absolute coordinates estimated by the principles of bundle adjustment. With a mean target distance of about 110 m, an average accuracy of ±1.8 cm was achieved. This demonstrates the potential of high resolution imagery for precise point estimation. The precision of the reconstruction of a scene or the estimation of flow velocity using optical imagery in general highly depends on the camera position geometry with respect to the object of interest. Also, optical distortion and other camera parameters need to be reliably estimated to accurately estimate valuable parameters from image data. In a recent study by *Kaufmann and Seier* (2016), the principles of bundle adjustment with a high number of ground control points (31) and structure from motion (SfM) techniques were used for scene reconstruction. Respective accuracies of about ±35 cm for position and ±6 cm for the height components, and reprojection errors in the order of 0.3 pixel were reported. Data acquisition was carried out once a year with four pre-defined camera positions.

For the purpose of prediction and/or filtering of time series, the method of collocation is frequently used in the geodetic context (e.g., *Hollenstein et al.*, 2008b; *Müller*, 2011; *Hurter*, 2014). As shown in *Geiger* (1996), the principles of collocation are comparable to the methods known as *kriging*, introduced by *Krige* (1951); *Matheron* (1963). Estimated or assumed correlations between measured quantities, linked by a geometrical distance (e.g., estimated coordinates of an object in space and time), are used to extract significant signal from noisy datasets. In scenarios of block-like displacement fields, an improved variant uses an iterative procedure to change the metric distance between the measurements. Specially in the vicinity of neighboring blocks, an improved characterization of the displacement field can be obtained. The concept of this adaptive collocation was introduced by *Egli* (2004); *Egli et al.* (2007). Similar approaches were made in the field of *kriging* (e.g., *Moustapha et al.*, 2015).

As stated above, the *X-Sense* project aims for high resolution monitoring in space and time. A first challenge, solved by the Computer Engineering Group, ETH Zürich, was to develop and build a reliable long-term infrastructure, with sensors that could be set up, measure, and survive in the alpine environment (*Beutel et al.*, 2011). The various measurement systems are embedded in a sensor network and allow near real-time data access. In this work, the principles of terrestrial photogrammetry, with a sparse network of two GNSS control points, are used to retrieve high resolution rock boulder motion estimates within the area of interest. The challenge is to get various 3D position time series with centimeter accuracy for the given difficult geometrical setting. The permanently installed GNSS and camera stations are operated throughout the year. Daily to sub-daily image intervals are used to verify the possible detection of acceleration events. With an average displacement rate of approximately 0.6 m/year (*Wirz et al.*, 2016) the challenge is to achieve an accuracy suitable to detect inter annual velocity variations for this type of slow permafrost creep.

## 1.4 Objectives and Structure of the Thesis

This study focuses on the combination of optical imagery and GNSS-based position estimates for the reconstruction of permafrost surface motion over a period of multiple years.

Within the *X-Sense* project, a stereo-pair of commercial DSLR (digital single reflex) cameras were set up in the vicinity of the observation area (Section 2.1). Two GNSS stations in field of view (FOV) measure the motion of two selected rock boulders. The goal of this work is to use this RAW data for the estimation of a number of high resolution 3D rock boulder trajectories. To reach this goal, four main objectives can be formulated:

- Development of a suitable feature tracking procedure (algorithms and strategy). The goal here is to have a flexible and highly accurate tool to estimate the image coordinates of individual features over short and long time intervals.
- Maximization of the usable time window for estimating the surface displacement rates. The high altitudes and orientation of permafrost areas also imply a relatively long winter period. The goal here is to also use partially snow covered images as valuable data sources.
- Optimal scene reconstruction by either using one camera and a digital elevation model or by combining the stereo-pair cameras with the two GNSS stations that are used as ground control points. The task of automatically combining the photogrammetric procedure with the available GNSS position estimates has to be addressed.
- Robust velocity estimation using the estimated feature positions. Because velocity is generally defined as the change in position within a given time interval (or the derivative of position w.r.t. time), noise in estimated feature trajectories has a large impact on velocity estimation. Therefore, a suitable method is needed to optimally filter the time series to improve the estimation of velocity.

The structure of the thesis closely follows the main objectives given above. Chapter 2 gives an overview of the study area, the equipment, and the processing strategy used. In Chapter 3, a detailed study and analysis of the least-squares matching technique, optimized for the problem statement, is given. Section 3.1 starts with an overview of related work that was conducted in the field of photgrammetry, specially in prospect of the least-squares matching principles. In Section 3.2, the mathematical principles of the technique are described. It comprises the estimation of initial parameters, outlier detection, robust estimators, and a performance test using synthetic examples. Section 3.3 then summarizes the data processing strategy used in the current study.

Chapter 4 covers the topic of object point reconstruction using one or more cameras. In Section 4.1, the coordinate system and conventions are defined. Section 4.2 explains the principles of 3D point estimation combining a single camera with a digital elevation model, emphasizing the problematic of error propagation in such cases. Section 4.3 and 4.4 then cover the mathematical principles of scene reconstruction using two or more cameras, the topic of parameter significance testing, gross error detection, and model errors. A simulation of an idealized scenario then addresses the question of limitation and accuracies that can be expected with this technique. The strategy to estimate initial parameters needed for this non-linear adjustment is given in Section 4.5. Section 4.6 finally illustrates the procedure being followed to combine the GNSS and image-based observations.

The principles of collocation, as a powerful tool for data filtering, prediction, and velocity estimation, are presented in Chapter 5. Introducing the principles of the technique in Section 5.1, the topic of estimating correlations among measurements is addressed in Section 5.2. The theoretical background of the adaptive collocation technique is given in Section 5.3. In Section 5.4, the derivations of the variance-covariance matrices for the respective quantities are formulated.

Having covered the theoretical part, key results are shown in Chapter 6. A time sequence of areal displacement rates estimated when using only a single camera is presented in Section 6.1. In Section 6.2, the achieved accuracies of the 3D point reconstruction along with examples of estimated time series of selected rock boulders are highlighted. Section 6.3 presents the collocated results obtained by filtering the noisy time sequences and also gives examples of the derived velocities. An example of the adaptive collocation technique applied to areal velocity fields is finally given in Section 6.4.

Chapter 7 discusses and summarizes the main objectives and findings of this thesis, and Chapter 8 gives a short outlook for possible future work in this research field.

# 2 Study Site

This chapter describes the study site relevant for this thesis. Within the *X-Sense* project, investigations at different permafrost landforms where conducted. For this task, GNSS measurement devices, high-resolution cameras, weather stations, crack-meters, and surface temperature loggers (iBottons, (*Gubler et al.*, 2011)) were deployed in the field. All devices are located within an area of approximately 6 km × 1.5 km and are located in the Mattervally, Valais, Switzerland. A more detailed presentation of the project study site and permafrost creep rates can be found in *Wirz et al.* (2016). The range of mean annual velocities, measured within the years 2011 to 2013, is between 0.006 and 6.3 m/year (GNSS-based). The focus of this thesis is in the upper Grabengufer permafrost creep area with mean annual velocities in the order of 0.6 m/year (not to mistake with the Grabengufer rock glacier at much faster speeds). Fig. 2.1 shows the location of the Grabengufer rock glacier and its associated rock slide area (*Delaloye et al.*, 2014).
Section 2.1 gives a short overview of the permafrost creep at the Grabengufer, Section 2.2 and Section 2.3 present the installations and RAW data that were deployed and used, and Section 3.3 explains the processing strategy applied in this study.

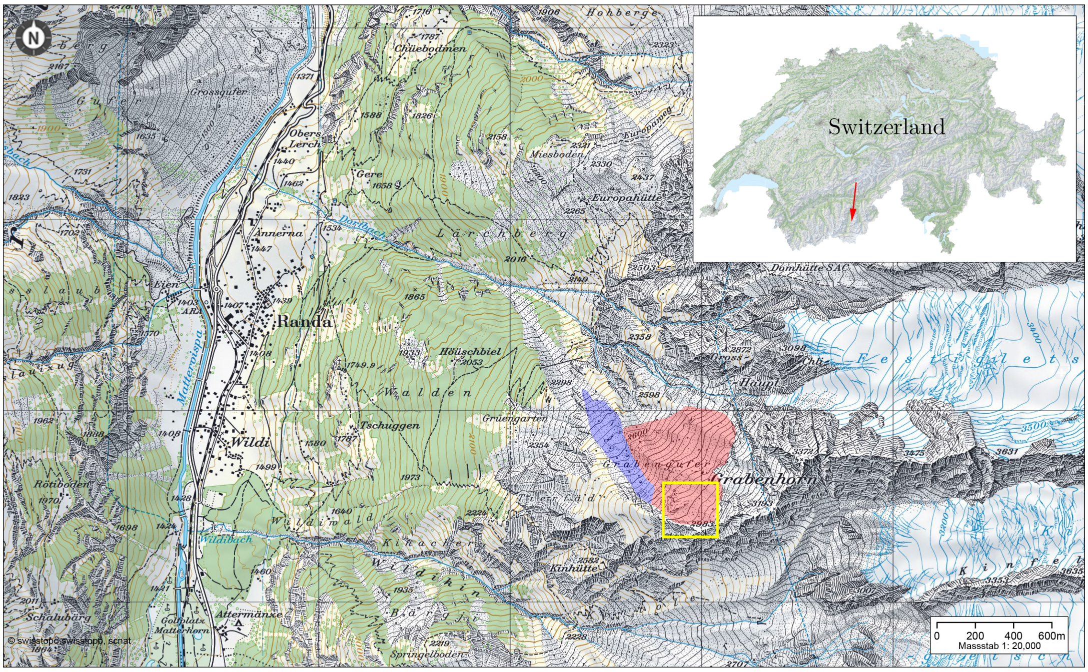


**Figure 2.1:** Location of the study site. The highlighted area in blue shows the Grabengufer rock glacier while the red area shows the region classified as the Grabengufer rock slide (*Delaloye et al.*, 2014). The yellow box on the southern part (uphill) indicates the area shown in Fig. 2.3. Topographic map LK 1:25'000, Swisstopo.

## 2.1 Grabengufer Permafrost Creep

Rock glaciers are geomorphological landforms located on periglacial mountain slopes. Typically they are composed of rock debris frozen in interstitial ice or they can be described as former ice glaciers that are now covered by a layer of talus. *Häberli* (1985) describes rock glaciers as «perennially frozen debris masses that creep down mountain slopes», already recognized by *Jäckli* (1957). This permafrost creep is a steady-state deformation of such debris masses that are supersaturated with ice. In 1985 about 1'000 active rock glaciers in Switzerland were known, whereas by the year 2010, about twice this number was estimated (e.g., *Delaloye et al.*, 2010). Active rock glaciers act as sediment conveyors with typical velocities of 0.1 to 2 m/year over long time scales of years to even centuries (*Delaloye et al.*, 2010). Studies using InSAR, (e.g., *Strozzi et al.*, 2009), have recently revealed several rock glaciers in the Mattervalley moving at high rates of more than 1 m/year. The Grabengufer rock glacier (Fig. 2.2) is among these so called destabilized rock glaciers with exceptional high displacement rates of up to 150 m/year, observed by *Delaloye et al.* (2013) during the period 2009 – 2010 (up to 40 cm/day). After this period, the displacement rates decreased. Destabilization processes of rock glacier tongues are assumed to be a response of warming ice and its consequences in the changing strain-stress relation (*Roer et al.*, 2008). The destabilization event caused rock fall and debris flow events further down the Grabengufer gully (*Bühler and Graf*, 2013).

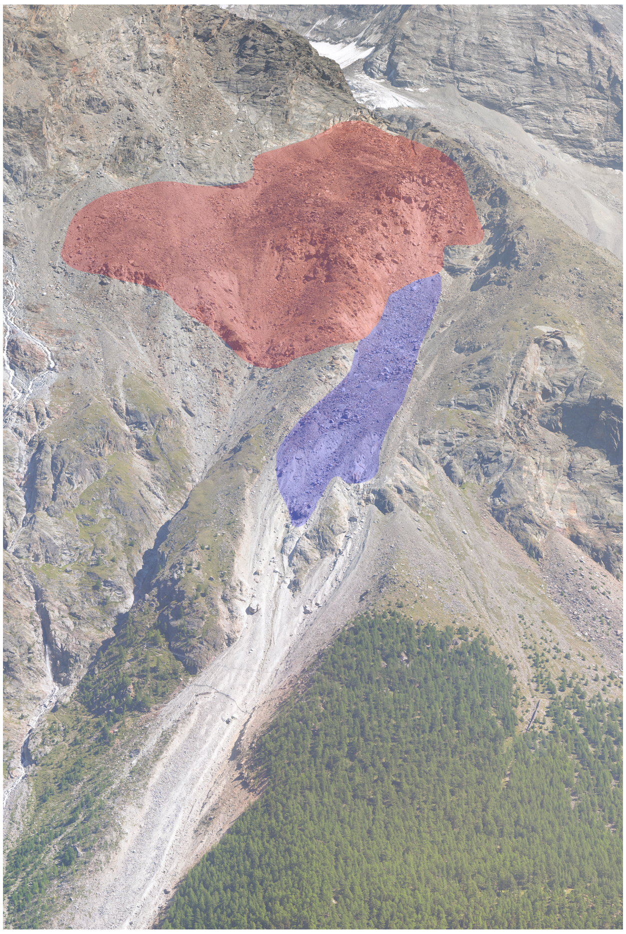

**Figure 2.2:** Grabengufer permafrost areas in August 2012. The approximate extent of the rock slide and rock glacier are given in red and blue, respectively. Background image data from Jan Beutel, ETH Zürich (*Keller et al.*, 2009).

As noted in *Delaloye et al.* (2013), the destabilization event of 2009 – 2010 is qualified as a mechanical surge process that started at the root zone of the rock glacier. In *Delaloye et al.* (2014), the area above the Grabengufer rock glacier is classified as the Grabengufer sag or rock slide (red area in Fig. 2.2), with velocities between 0.1 and 0.8 m/year between the years 2012 and 2013. The area just above the rock glacier was observed as the origin of several rock falls with volumes of up to 4300 $m^3$ (observed in 2010, also destroying the hanging bridge about 600 m further downslope). For this study, the area just above this tear-off edge is monitored.

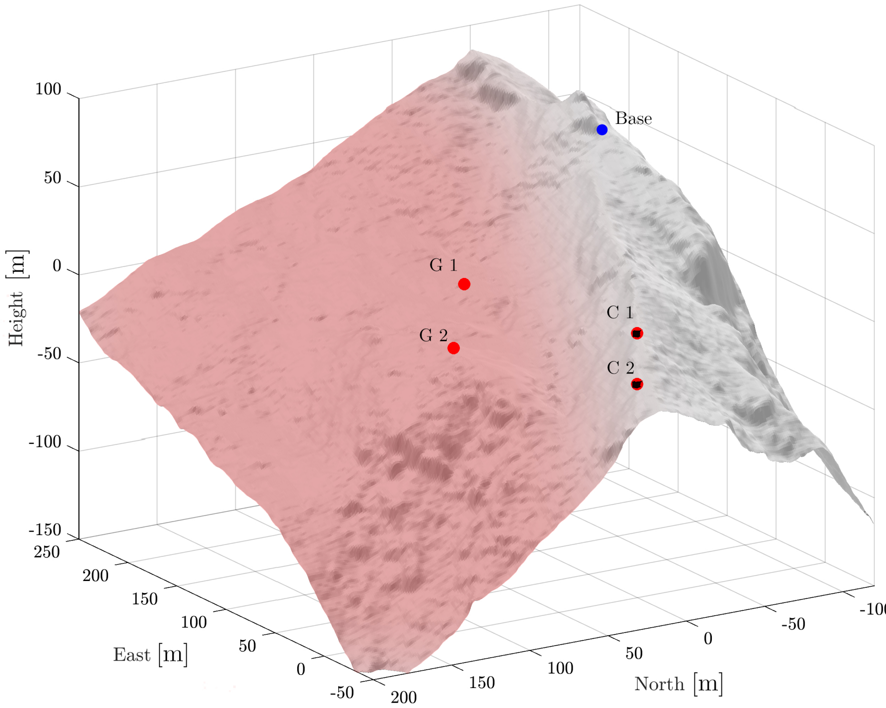


**Figure 2.3:** Hardware deployments at the field site (the rock slide zone is shaded in red). G1 and G2 mark the moving GNSS stations, C1 and C2 the two camera stations that are equipped with GNSS modules. A base station on a solid rock is used as local reference position. The coordinate system is given in a topocentric coordinate frame (see Section 4.1).

## 2.2 Field Installations

Measurement devices were installed in different areas of the project study site. The design, development, deployment, and operation of these systems (i.e., hard- and software for GNSS and camera stations) were conducted by the Computer Engineering Group, ETH Zürich. The individual components were build to sustain the harsh conditions in the alpine environment (*Beutel et al.*, 2009).

Within the Grabengufer rock slide (Section 2.1), two prototype low-cost GNSS devices suitable for the extreme alpine environment were deployed as high-precision measurement systems to monitor the position of two selected rock boulders through time[1]. The measurement network at the Grabengufer site consists of five GNSS measurement positions and two optical camera systems. Fig. 2.3 shows the distribution of the main installations, with individual components described as:

- Base station with GNSS receivers and communication modules
- Two low-cost GNSS stations, G1 and G2, mounted on large rock boulders
- Two commercial digital single reflex cameras, including GNSS modules, C1 and C2

[1] One station (G2) is a dual GNSS device, i.e., two antennas are mounted on the same mast. For the purpose of this study, only the position estimates of the antenna centered over the mast are used.

Additionally, a crack-meter near G2, a weather station at the base station, and surface temperature loggers were deployed in the course of the *X-Sense* project but not used in this study. Data packages acquired at each station are transfered to the base station by a multi-hop procedure using a local wireless network (see *Beutel et al.* (2011) for more details). A connection of the local wireless sensor network to the Internet is established at the base station (and at the two camera stations), where generally more power is available. All data (GPS data, camera images, meteorological data, etc.) is stored on a server in Zürich (Computer Engineering Group, ETH Zürich), allowing near real-time data access.

### 2.2.1 GNSS Stations

The permanent GNSS stations consist of different components: A GNSS module (antenna and L1-GPS frequency receiver - ublox LEA-6T), a two-axis inclinometer (SCA830-D07), a photo-voltaic energy harvesting system including a battery acting as a buffer, wireless radios, and a mast (1 m in height) where all the components are mounted to (Fig. 2.4). These prototype low energy consumption devices provide the ability to send and receive data packages, also acting as hop stations for other devices. A detailed describtion of the principle components and methodology can be found in *Wirz et al.* (2013) and a more detailed describtion of the prototype components is given in *Buchli et al.* (2012). Data logging follows a two-stage power plan: In case the battery voltage is high enough, con-

**Figure 2.4:** One of the two *X-Sense* low-cost GNSS stations mounted on a large boulder within the Grabengufer rock slide area. Image Jan Beutel, ETH Zürich (*Buchli et al.*, 2012).

tinuous measurements are carried out (typically at a sampling interval of 30 seconds). If the battery voltage drops below a certain threshold (typically 11.8 V for a 12 V battery), a low power operation mode is activated. During this mode, only sparse measurements (e.g., from 11:00 to 13:00 CEST) are carried out. The latter can be tuned and therefore allows to maximize the valuable output given the energy available (more details in *Buchli et al.* (2014)).

Static daily GNSS solutions with cm to mm accuracies are computed automatically using the base station as a local reference and Bernese processing software (*Beutler et al.*, 2007; *Limpach and Grimm*, 2009). The principle GPS processing strategy is based on single-frequency differential carrier phase techniques and is used for all GNSS stations that are integrated into the *X-Sense* project (implemented and operated at the group of Mathematical and Physical Geodesy, ETH Zürich).

Rock boulders are exposed to local rotations such that variations in estimated GNSS coordinates, representative for a specific boulder, may be misinterpreted as pure translation. *Wirz et al.* (2014) suggested a possible workflow to combine orientated inclinometer measurements with antenna position estimates. Along with some assumptions, the true translation at the GNSS mast base can then be derived. In this thesis, however, GNSS solutions are used in a different way, i.e., the GNSS antennas are directly localized in the images, such that these effects are not of concern.

### 2.2.2 Camera Stations

A stereo pair of Nikon D300s cameras, equipped with 14 mm focal length wide angle lenses, were installed next to the study site in October 2012. With a mean elevation of about 2'900 m and a location exposed to strong winds in the harsh alpine environment, a massive and imperishable construction was built (C. Senn, D-BAUG, ETH Zürich) to mount the camera box in the field (Fig. 2.5). Solar panels charge a 12 V battery that powers the camera, the GNSS module, and the WLAN antenna for data transmission. The latter allows to access the images in almost real time. An embedded PC platform is running Linux that enables communication with the camera to allow, e.g., to change the frequency of image acquisition. A more detailed description of this camera system can be found in *Keller et al.* (2009). The system is embedded in the *X-Sense* sensor network such that images are automatically downloaded to the server at ETH in Zürich.

Camera motion is monitored by the on-board GNSS module and daily solutions are generated using the same principles as for all other GNSS stations within *X-Sense*. A maximum displacement of approximately 8 cm is observed for camera 1 (C1) in the East component over the course of 3.5 years (details are in Fig. E.1). This station is affected by periodic displacements, i.e., during the warm season displacements in the range of 2 − 3 cm are observed, whereas it remains stable during the cold period. The motion of camera 2 (C2) is different: a continuous displacement with a temporary acceleration phase that happened during summer 2015 was measured. Over the course of 3.5 years, the total absolute displacements are 10.8 cm and 6.2 cm for station C1 and C2, respectively.

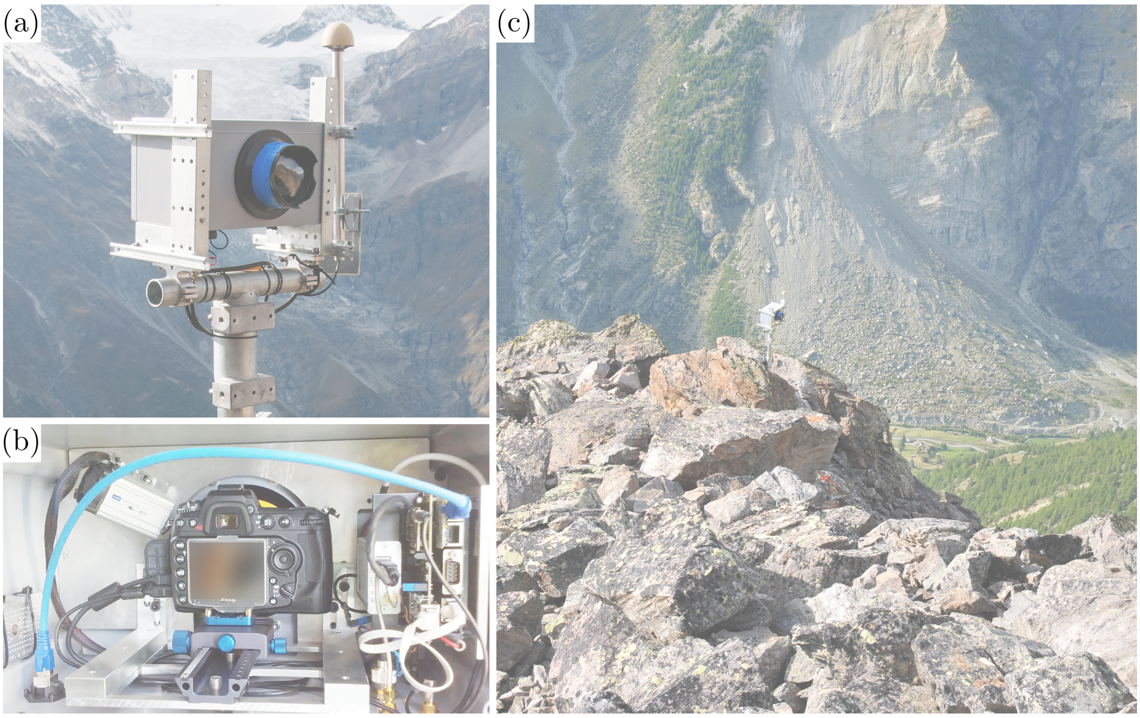

**Figure 2.5:** One of the two camera stations (C2) installed at the study site. (a) shows a closeup of the massive steel construction along with the GNSS antenna mounted next to the camera box. The interior of the camera installation is shown in (b), image Tonio Gsell, ETH Zürich, with details in *Keller et al.* (2009). An impression of the terrain surrounding the installation site is given in (c). The Randa rock fall that happened in 1991 is visible in the background.

## 2.3 Measurement Period and Data

Data collection at the Grabengufer field site started in the year 2011 (GNSS stations) and is currently (by the end of 2016) still ongoing. As stated above, the two cameras were installed in October 2012, also representing the start of the measurements used for this work. While both the GNSS and the camera stations worked with little to no interruption throughout the years, only periods between (typically) June and November could be used. During the rest of the year, the area was fully covered by snow so that no photogrammetric solution could be computed. July 2016 marks the end date of the data used herein.

All results presented in this thesis are based on the position solutions of the two GNSS stations on the Grabengufer rock slide, the GNSS position solutions of the two camera installations, and the RAW images from the stereo-pair cameras between October 2012 and July 2016. While one solution per day is typically available for the GNSS results, images were usually captured in an hourly interval during daytime. Because of the slow permafrost motion, only one image per day (in average) was used for further analysis. RAW data for both, GPS signal measurements and the camera images were provided by the Computer Engineering Group, ETH Zürich. To test the monoplotting procedure, a digital elevation model (provided by FOEN) from the year 2010 was used in addition.

## 2.4 Processing Strategy

The goal of this study is to combine position estimates of low-cost single-frequency (L1) GNSS sensors with image sequences obtained by off-the-shelf DSLR cameras. Data merging should result in a precise estimate of the permafrost creep in various areas that are within the field of view. As mentioned in the previous chapter, different possibilities exist to tackle such problems. As soon as high precision is of interest, the method of least-squares-matching (LSM) is the optimal choice to identify the same features in different images (*Gruen*, 1985), more details in Section 3.1.

Absolute quantities of motion or position coordinates can be obtained by the combination of at least two cameras (stereo-/multi-view) or by the principles of ray tracing onto a Digital Elevation Model (DEM), known as monoplotting (e.g., *Gruen and Sauermann*, 1977). In the course of this work, both methods are studied. In a first attempt, one camera combined with a DEM is used to estimate the displacement rates. Scaling values for displacements in the image space to absolute units are obtained by ray tracing onto a DEM. The workflow for this task is given in Fig. 2.6.

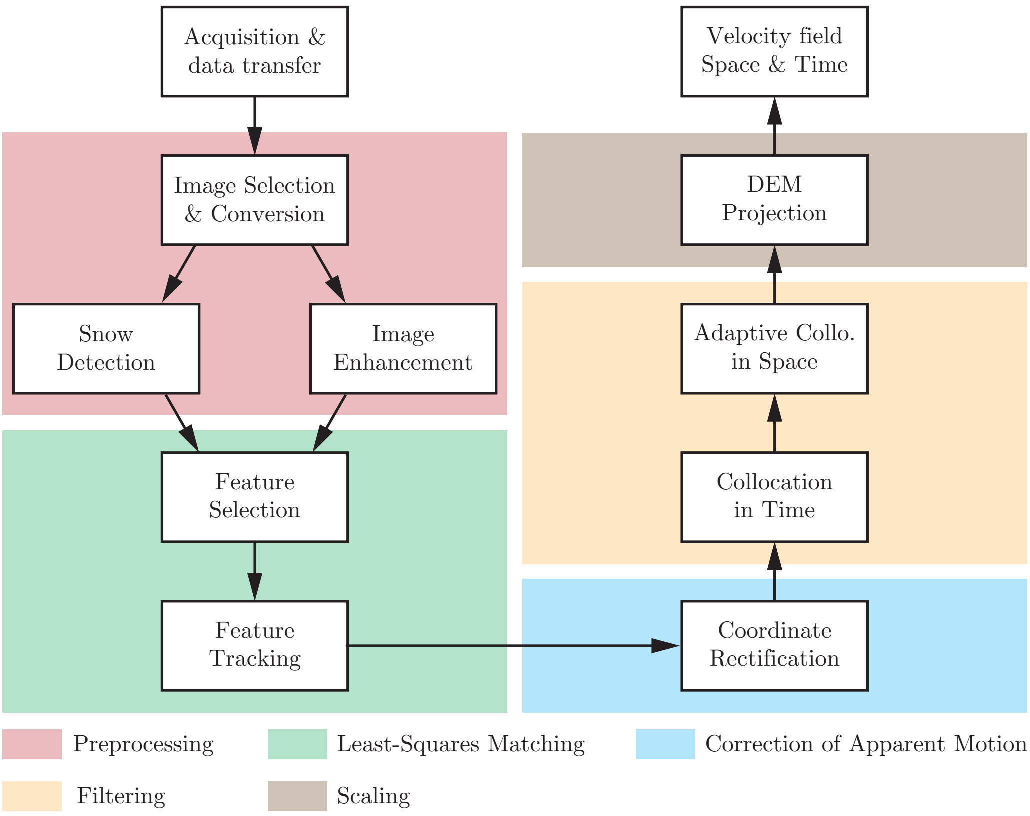


**Figure 2.6:** Processing flow chart of estimating velocity fields using a single camera and a DEM.

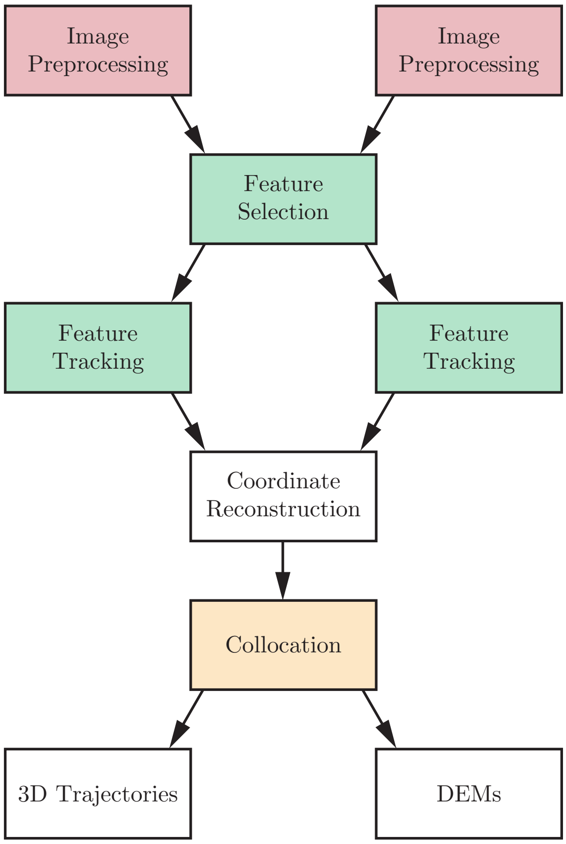


**Figure 2.7:** Multi-view processing flow chart for the reconstruction of object trajectories and DEMs. The colorization is according to Fig. 2.6. Coordinate reconstruction is accomplished using the principals of bundle adjustment.

The principle of bundle adjustment is used to reconstruct the scene for various features (see Fig. 2.7). Using the LSM technique to estimate feature coordinates in the image sequences of both cameras, an epoch-wise estimation of object coordinates yields the geo-referenced feature trajectories in space (3D). Errors in the image matching process are carried on into the bundle adjustment that itself is an error minimization process of observations. As in the case of all adjustment processes, observations are linked by a mathematical model that approximates reality to a certain extent. Model and other error sources lead to results that are contaminated by noise. Filtering these noisy time series is then realized by the method of collocation using estimated correlations among the position trajectories of individual features.

The following three chapters describe these three main tasks, namely least-squares image matching, object reconstruction in space, and collocation.

# 3 Image-Based Displacement Estimation

Optical imagery with off-the-shelf digital cameras is a popular measurement technique for a variety of applications. Although the principle is known for many decades, improved processing strategies and easy-to-use applications promote the popularity of this measurement method. A number of recent studies thus explicitly us image-based observations for monitoring purposes (e.g., *Travelletti et al.*, 2012; *Scaioni et al.*, 2014). While some use specific targets for detection and point localization (*Kersten and Mass*, 1995), natural targets or a combination of both are more frequently used nowadays (e.g., *Maas*, 1996). There are scenarios, where no artificial targets can be installed on the object of interest or where the expense of such installations would no longer be adequate (e.g., steep rock walls with continuous breakouts).

In this chapter, the principle of image-based displacement estimation based on natural targets is introduced. Section 3.1 gives an overview about possible mathematical approaches as well as the motivation to choose the adaptive Least-Squares Matching (LSM) solution for this type of study. In Section 3.2, the mathematical principle of the LSM method is outlined, along with statistical testing and gross error detection. Finally, in Section 3.3, the processing strategy is discussed.

## 3.1 Conceptual Overview

The estimation of motion from or between image sequences is a problem known for many years and different solutions exist for a variety of applications (motion estimation in surveillance (e.g., *Wang and Brady*, 1995), motion based segmentation (e.g., *Sturm and Triggs*, 1996; *Mitiche and Ayed*, 2011), structure and depth from motion (e.g., *Häming and Peters*, 2010), obstacle avoidance (e.g. *Nelson and Aloimonos*, 1989), image composition and registration (e.g., *Zitová and Flusser*, 2003)). Many implementations are used today in industry, performing daily tasks in medical image analysis, driver assistance, navigation of drones, etc. The principle in all these applications is to use the distribution and location of image pixel intensities to retrieve valuable information about the scene.

For monitoring ice- or rock glaciers and creeping phenomena in general, optical image sequences have previously been used either as terrestrial (e.g., *Travelletti et al.*, 2012) or air-/space-borne (e.g., *Kääb*, 2002) datasets. To estimate motion of such objects, sparse feature matching techniques are typically applied to find the position of corresponding features in images taken at different epochs. In contrast, *Vogel et al.* (2012) presented methods that perform dense flow estimation, i.e., a displacement vector is estimated for every image pixel. As this problem is ill-posed (known as the aperture problem), regularization methods in image space and time are applied. Although there are so called dense matching techniques that try to estimate large displacements in the local neighborhood, the principle relies on the assumption of smooth displacements between neighboring pixels and adjacent epochs. Specially in cases of a regular sampling (e.g. video sequences) and known smooth motion transitions, this a priori information can successfully be used to stabilize flow estimation. Estimating motion using these techniques is known as optical flow estimation. This field was extensively studied in the last decades and a number of strategies have evolved (*Barron et al.*, 1994). The algorithms range from classical methods (e.g., *Horn and Schunck*, 1981; *Lucas and Kanade*, 1981) to more recent developments like cost-volume filtering or modified versions of the former two ( *Vogel et al.*, 2012).

One of the challenges when monitoring permafrost creep over long periods by optical imagery, is to deal with the irregular sampling: periods of bad weather, sequences of snow coverage, specially in autumn, and the winter season, with only a few (if any) features suitable for matching. In the given geometrical setting of the study field (see Section 2.2), relatively fast moving areas project next to stable (or slowly moving) areas further in the background, yielding a sharp boundary of motion. As observed by *Delaloye et al.* (2010), sharp boundaries of motion may also exist within the flow field, individual rocks can show rotational components other than their neighbourhood and the kinematics of permafrost creep may show significant fluctuations in short time scales. As GNSS stations are optically monitored and integrated into the reconstruction process (Chapter 4), an adaptive feature matching technique (introduced by *Gruen* (1985) and extended in *Baltsavias* (1991)) is used to solve for the displacement of individual rocks (Section 3.2) as well as for matching the GNSS template (Section 4.6). The mathematical principle (described in Section 3.2.1) allows to correct for local rotations, scale variations as well as affine and perspective distortions.

In a general formulation of the Least-Squares Matching (LSM) technique, a combined solution of image feature tracking and object point reconstruction can be formulated ( *Gruen*, 1985; *Baltsavias*, 1991). Using this technique, procedures for particle tracking in the object space were successfully applied (*Maas et al.*, 1993).

For this study, the problem of 3D tracking is divided into two separate adjustment problems (LSM and scene reconstruction). The disadvantage by doing so is that the joint system of motion detection and object point reconstruction is lost. On the other hand, this method only works well if the camera parameters are known or if the mathematical model is extended to included all features of multiple epochs as image pixel observations (for example mentioned in *Triggs et al.* (2000)). In the latter case, connection in time is given by the feature transformation parameters, whereas in space it is defined by the camera parameters. In such a scenario, the observation and parameter spaces become extremely large. Because the LSM technique is also used for other matching tasks, the problem is divided into two separate processes.

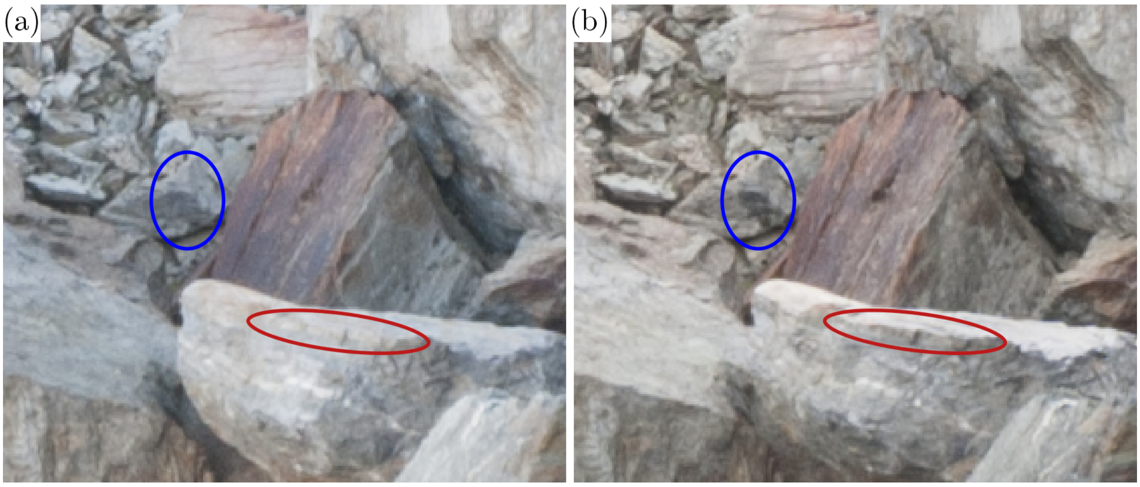


**Figure 3.1:** Effect of different scene reflectance and illumination. Image (a) was taken on 01-08-2014 08:00 UTC, image (b) on 06-08-2014, 08:00 UTC. Although the scene is illuminated by diffuse scattering in both cases, obvious differences exist. The red ellipse highlights a zone, where the edge of a rock bolder appears differently, the blue ellipse points to an area before and after water outflow, obviously changing the reflectivity.

### 3.1.1 Motivation and Benefits of Feature Tracking

Using a local LSM for feature tracking is suitable for a variety of different tasks. For example the same procedure can in principle be used for matching natural features between images, for matching a given template with respect to a specific pixel location (Section 4.6) or for matching structures across different views in order to reconstruct a scene (e.g., *Lai*, 2000).

An additional benefit of using LSM is that the statistical characterization of the solution is well established. Estimated feature transformations are uncorrelated among each other if the matching windows (Section 3.2.4) do not overlap, a condition that is not true in case the optical flow is computed using dense methods (*Fermüller et al.*, 2001). Thus, gross error detection, parameter significance testing, and uncorrelated parameter estimates between different feature displacements are the advantages.

### 3.1.2 Limitations and Expected Accuracy

Although the principle matching accuracy might be very high, the accuracy and reliability of the displacement estimation is heavily influenced by varying scene illumination, varying surface reflectivity (e.g., dry or wet surface), and occlusions (e.g., areas partly covered with snow). Even if the sun does not directly strike the scenery, differences in diffuse light scattering cause the reflected light to vary such that obvious differences in the appearance result (see Fig 3.1 for an example).

Due to the large variability of surface structures, similarities of gradient structures in rock agglomerations, and the variable illumination condition between two images, false positive matching results might occur even if the algorithm is outlined for high robustness. For example *Ackermann* (1984) and *Danuser* (1996) report a maximum matching accuracy for the translation components of 0.05 pixel, obtained under ideal conditions. In case of long-term outdoor image sequences with changing environmental conditions, the accuracy limit is assumed to be in the range of 0.1 to 0.5 pixel.

## 3.2 Least-Squares Feature Tracking

The principle of Least-Squares Matching (LSM) is a well known and a popular method for various applications. *Gruen* (1984, 1985), *Ackermann* (1984), and *Pertl* (1984) developed in parallel the mathematical principals of this technique. Due to its flexibility, different problems can be solved using the same implementation: single view image feature matching (e.g., *Zhang and Gruen*, 2006), wide baseline matching techniques for stereo or multi-view 3D object point reconstruction (e.g., *Brown et al.*, 2003; *Tuytelaars and Van Gool*, 2004; *Zhang and Gruen*, 2006), and other photogrammetric matching problems (e.g., *Gruen and Akca*, 2005; *Beyer*, 1992). Although there are some disciplines, where more efficient algorithms exist (*Brown et al.*, 2003; *Baumberg*, 2000), the principle of LSM still remains a powerful, flexible, multi-purpose, and very accurate method to solve a variety of photogrammetric problems.

Sections 3.2.1-3.2.3 describe the mathematical and statistical model of the LSM technique, mostly based on the work of *Gruen* (1985) and *Baltsavias* (1991). Sections 3.2.4 and 3.2.5 discuss methods regarding the general estimation procedure along with initial parameter estimations, Section 3.2.6 introduces robust estimators, and Section 3.2.7 gives an overview of the expected performance using simulated examples.

### 3.2.1 Mathematical Model

A least-squares problem in its very basic form can be written as a relation between some observations $\mathbf{l}$ and a model $F$ with parameters $\mathbf{p}$:

$$F(\mathbf{p}, \mathbf{l}) = 0 \tag{3.1}$$

The actual observations are always subject to errors ($\mathbf{l} = \check{\mathbf{l}} + \check{\mathbf{e}}$), whereas their expectation values $E\langle\mathbf{l}\rangle$ are $\check{\mathbf{l}}$. As the true observation and error components cannot be known, a best estimate of those have to be used: $\hat{\mathbf{l}}$ and $\hat{\mathbf{e}}$. The very principle of the image LSM algorithm is to estimate a set of parameters $\hat{\mathbf{p}}$ such that the sum of squared differences between the first image (further called the 'template' image, symbol $\mathbf{h}$) and the transformed second image (further called the 'patch' image, symbol $\mathbf{g}$) is minimal. Regarding the general formulation (Eqn. (3.1)), the image matching problem can be formulated as as:

$$\mathbf{h}(\hat{\mathbf{l}}_h + \hat{\mathbf{e}}_h) - \mathbf{g}(\hat{\mathbf{p}}, \hat{\mathbf{l}}_g + \hat{\mathbf{e}}_g) = 0 \tag{3.2}$$

Provided that the measurement errors are normally distributed and assuming that $\mathbf{h}$ and $\mathbf{g}$ are linear in $\hat{\mathbf{l}}_h$ and $\hat{\mathbf{l}}_g$, respectively, the least-squares parameter estimator can be approximated by a maximum likelihood estimator (MLE) (e.g. *Danuser and Stricker*, 1998). With these assumptions, the formulation above can be written as a Gauss-Markov estimation model:

$$\mathbf{h}(\hat{\mathbf{l}}_h) + \hat{\mathbf{e}}_h - \mathbf{g}(\hat{p}, \hat{\mathbf{l}}_g) - \hat{\mathbf{e}}_g = 0 \tag{3.3}$$

$$\mathbf{h}(\hat{\mathbf{l}}_h) - \mathbf{g}(\hat{p}, \hat{\mathbf{l}}_g) + \hat{\mathbf{v}} = 0 \tag{3.4}$$

where $\hat{\mathbf{v}}$ equals the difference of both residuals, i.e., $\hat{\mathbf{v}} = \hat{\mathbf{e}}_h - \hat{\mathbf{e}}_g$. The first term in Eqn. (3.4) is the template image, i.e., an estimate of the error-free observations, in the coordinate system $[x, y]$. Consequently, the second term describes the modeled patch image (with the parameter set $\hat{\mathbf{p}}$ in the coordinate system $[u, v]$. Eqn. (3.4) can also be written as:

$$\hat{\mathbf{h}}[x, y] + \hat{\mathbf{v}}[x, y] = \mathbf{F}(\hat{p}_1, .., \hat{p}_N, \hat{\mathbf{g}}[u, v]) = \tilde{\mathbf{g}}[x, y] \tag{3.5}$$

$$\mathbf{v} \sim \mathcal{N}(0; \sigma_0^2 \cdot \mathbf{Q}_{hh}) \tag{3.6}$$

with $\hat{\mathbf{h}}[x, y]$, $\hat{\mathbf{g}}[u, v]$ being the template and the patch observations in their respective coordinate systems, $\mathbf{F}(..)$ the transformation of the patch image (from the coordinate system $[u, v]$ to the coordinate system $[x, y]$ of the template image) with estimated parameters $\hat{p}_1, .., \hat{p}_N$, $\tilde{\mathbf{g}}[x, y]$ the transformed and resampled patch image, and $\mathbf{Q}_{hh}$ as the cofactor, matrix. Resampling onto a common grid is necessary for evaluating the similarity (residual) component.

To solve the non-linear Gauss-Markov model, the parameters are determined iteratively using the linearized expression with initial estimates $\mathring{\mathbf{p}}$ as:

$$\begin{aligned} \hat{\mathbf{h}} + \hat{\mathbf{v}} &= \mathbf{F}(\hat{p}_1, .., \hat{p}_N, \hat{\mathbf{g}}) \\ &= \tilde{\mathbf{g}}(\mathring{p}_1, .., \mathring{p}_N, \mathbf{g}) + \left.\frac{\partial \tilde{\mathbf{g}}}{\partial \mathbf{p}}\right|_{\circ} \cdot (\hat{p} - \mathring{p}) = \mathring{\mathbf{g}} + \mathbf{A} \cdot \delta\hat{\mathbf{p}} \end{aligned} \tag{3.7}$$

with

$$\mathbf{A} = \left[ \left.\frac{\partial \tilde{\mathbf{g}}}{\partial p_1}\right|_{\circ} \left.\frac{\partial \tilde{\mathbf{g}}}{\partial p_2}\right|_{\circ} \cdots \left.\frac{\partial \tilde{\mathbf{g}}}{\partial p_N}\right|_{\circ} \right] \tag{3.8}$$

$$\delta\hat{\mathbf{p}} = \left(\mathbf{A}^T \mathbf{Q}_{hh}^{-1} \mathbf{A}\right)^{-1} \cdot \mathbf{A}^T \mathbf{Q}_{hh}^{-1} \delta\hat{\mathbf{h}} \tag{3.9}$$

$$\delta\hat{\mathbf{h}} = \hat{\mathbf{h}} - \mathring{\mathbf{g}} \tag{3.10}$$

$$\mathbf{Q}_{hh} = \text{cofactor matrix of the observation vector } \mathbf{h} \tag{3.11}$$

$$\hat{\mathbf{p}} = \mathring{\mathbf{p}} + \delta\hat{\mathbf{p}} \tag{3.12}$$

During each iteration, the variables $\hat{\mathbf{p}}$, $\mathring{\mathbf{g}}$, $\mathbf{A}$, and $\delta\mathbf{h}$ are updated. The iteration stops, if the update of all the parameters being determined fall below a certain threshold. Regarding the Jacobian matrix $\mathbf{A}$, the partial derivative for the $k$-th observation equation and the $i$-th parameter is formed as:

$$\left.\frac{\partial \tilde{\mathbf{g}}^k}{\partial p_i}\right|_{\circ} = \frac{\partial \tilde{\mathbf{g}}_{\circ}^k}{\partial x_k} \cdot \frac{\partial x_k}{\partial p_i} + \frac{\partial \tilde{\mathbf{g}}_{\circ}^k}{\partial y_k} \cdot \frac{\partial y_k}{\partial p_i} \tag{3.13}$$

The terms $\frac{\partial \tilde{\mathbf{g}}_\circ^k}{\partial x_k}$ and $\frac{\partial \tilde{\mathbf{g}}_\circ^k}{\partial y_k}$ are the respective $x$ and $y$ gradients of the transformed patch image $\tilde{\mathbf{g}}[x,y]$ at position $k$. $\frac{\partial x_k}{\partial p_i}$ and $\frac{\partial x_k}{\partial p_i}$ are the partial derivatives of the transformed coordinates $\mathbf{g}[u,v] \to \mathbf{g}[x,y]$, represented by the coordinate transformation matrix $\mathbf{T}$. For an affine transformation, using the notation of homogeneous coordinates, this reads as:

$$\begin{pmatrix} x & y & 1 \end{pmatrix}_k = \begin{pmatrix} u & v & 1 \end{pmatrix}_k \cdot \begin{pmatrix} p_1 & p_2 & 0 \\ p_3 & p_4 & 0 \\ p_5 & p_6 & 1 \end{pmatrix} = \begin{pmatrix} p_1 \cdot u + p_3 \cdot v + p_5 \\ p_2 \cdot u + p_4 \cdot v + p_6 \\ 1 \end{pmatrix}_k^T \tag{3.14}$$

and thus the partial derivatives with respect to the six parameters $p_1, \ldots, p_6$ are:

$$\begin{aligned} \frac{\partial x_k}{\partial p_1} &= u_k & \frac{\partial y_k}{\partial p_1} &= 0 \\ \frac{\partial x_k}{\partial p_2} &= 0 & \frac{\partial y_k}{\partial p_2} &= u_k \\ \frac{\partial x_k}{\partial p_3} &= v_k & \frac{\partial y_k}{\partial p_3} &= 0 \\ \frac{\partial x_k}{\partial p_4} &= 0 & \frac{\partial y_k}{\partial p_4} &= v_k \\ \frac{\partial x_k}{\partial p_5} &= 1 & \frac{\partial y_k}{\partial p_5} &= 0 \\ \frac{\partial x_k}{\partial p_6} &= 0 & \frac{\partial y_k}{\partial p_6} &= 1 \end{aligned} \tag{3.15}$$

Adding two radiometric correction parameters (contrast and offset), the matrix $\mathbf{A}$ for an affine transformation defined by Eqn. (3.14) becomes:

$$\mathbf{A} = \begin{pmatrix} g_{x,1} \cdot u_1 & g_{y,1} \cdot u_1 & g_{x,1} \cdot v_1 & g_{y,1} \cdot v_1 & g_{x,1} & g_{y,1} & \mathring{g}_1 & 1 \\ g_{x,2} \cdot u_2 & g_{y,2} \cdot u_2 & g_{x,2} \cdot v_2 & g_{y,2} \cdot v_2 & g_{x,2} & g_{y,2} & \mathring{g}_2 & 1 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ g_{x,k} \cdot u_k & g_{y,k} \cdot u_k & g_{x,k} \cdot v_k & g_{y,k} \cdot v_k & g_{x,k} & g_{y,k} & \mathring{g}_k & 1 \end{pmatrix} \tag{3.16}$$

The image gradients in $x$ and $y$ directions ($\mathbf{g}_x$ and $\mathbf{g}_y$) have to be approximated from the discrete image $\tilde{\mathbf{g}}[x,y]$, as there is no analytical form. As suggested in *Farid and Simoncelli* (2004), the estimate of the image gradient is computed by using a two-dimensional convolution filter with predefined convolution kernels. In case of excessive image noise, these convolution kernels can be modified such that an edge softening effect results, generally stabilizing the adjustment procedure. Here, two optimal 5-element vectors are used.

An estimate of the parameter accuracy is given by the parameter variance-covariance matrix $\mathbf{K}_{\hat{p}\hat{p}}$:

$$\mathbf{K}_{\hat{p}\hat{p}} = \hat{\sigma}_0^2 \cdot \left(\mathbf{A}^T \mathbf{Q}_{hh}^{-1} \mathbf{A}\right)^{-1} = \hat{\sigma}_0^2 \cdot \mathbf{Q}_{\hat{p}\hat{p}} \tag{3.17}$$

The choice of the observation cofactor matrix $\mathbf{Q}_{hh}$ is usually relatively simple. Because in general no information is given about the pixel variances, a unique variance (typically 1) is assigned to all observations without any correlations between them. With mean values $\mu_{h,g}$, the observations of the template and patch images $\mathbf{h}$, and $\mathbf{g}$ are generally assumed to be independent and normally distributed, with variance $\sigma_0^2$:

$$\mathbf{h}, \mathbf{g} \sim \mathcal{N}(\mu_{h,g}; \sigma_0^2 I) \tag{3.18}$$

This might be true for the error-free image, but there are typically four effects that violated this relation (Eqn. (3.18)):

(1) Light scattering

(2) Bayer demosaicing

(3) Image resampling

(4) Gaussian smoothing

As light between the object and the camera sensor travels through different media (various air layers and optical interfaces), light scattering effects (1) are present all the time and generally increase with increasing travel distance. Adding the natural air turbulance results in blurred projections of the objects. If this effect does not remain below the pixel resolution, the light of a point object will be spread across more than a single pixel, thus these observations are correlated. When using amateur DSLR cameras, every image is produced by a debayering process (2). Because every pixel in these single shot images record only one of the primary colors (Red, Green, or Blue) the missing colors have to be estimated by interpolation across the local pixel neighborhood (*Keigo Hirakawa*, 2005; *Ramanath et al.*, 2002). A possibility to avoid this interpolation step is to create a single pixel out of a 2x2 pixel patch containing all the color information. The latter process, however, reduces the resolution and might thus only be considered under special circumstances (for example in szenarios, where the pixel resolution is larger than the theoretical resolution of the optical components).

During the adjustment of the non-linear system (Eqn. (3.5)), another pixel-to-pixel correlation is introduced by the image resampling process (3) that typically consists of a bilinear or a bicubic interpolation. In case the parameter estimation process shows a behavior of alternating parameter convergence or if the image has a bad signal-to-noise ratio, Gaussian smoothing (4) is commonly applied to increase stability (*Berger*, 1999), introducing high correlations between neighboring pixels. If these correlations are neglected, the estimated parameters tend to have a variance that is too optimistic, even if the parameter estimation process is not significantly influenced. Mathematically, the effect on the covariance matrix $\mathbf{Q}_{\hat{p}\hat{p}}$ of the parameters can be assessed by analyzing the effect of an error $\mathbf{\Delta P} = (\Delta \mathbf{Q}_{hh}^{-1}$ in the weight matrix $\mathbf{P}^* \rightarrow \mathbf{P} + \Delta\mathbf{P}$ (see *Koch* (1988) for more details):

$$\begin{aligned} \delta\hat{\mathbf{p}} &= \left(\mathbf{A}^T\mathbf{P}\mathbf{A}\right)^{-1} \cdot \mathbf{A}^T\mathbf{P}\delta\mathbf{h} \\ \delta\hat{\mathbf{p}} + \Delta\delta\hat{\mathbf{p}} &= \left(\mathbf{A}^T\mathbf{P}\mathbf{A} + \mathbf{A}^T\Delta\mathbf{P}\mathbf{A}\right)^{-1} \cdot \left(\mathbf{A}^T\mathbf{P}\delta\mathbf{h} + \mathbf{A}^T\Delta\mathbf{P}\delta\mathbf{h}\right) \end{aligned} \tag{3.19}$$

Assuming small elements in $\Delta\mathbf{P}$, thus neglecting products with $\Delta\mathbf{P} \cdot \Delta\mathbf{P}$, this can be written as:

$$\delta\hat{\mathbf{p}} + \Delta\delta\hat{\mathbf{p}} = \left(\mathbf{A}^T\mathbf{P}\mathbf{A}\right)^{-1} \cdot \left(\mathbf{A}^T\mathbf{P}\delta\mathbf{h} - \mathbf{A}^T\Delta\mathbf{P}\mathbf{A}\left(\mathbf{A}^T\mathbf{P}\mathbf{A}\right)^{-1} \cdot \mathbf{A}^T\mathbf{P}\delta\mathbf{h} + \mathbf{A}^T\Delta\mathbf{P}\delta\mathbf{h}\right) \tag{3.20}$$

**Table 3.1:** Comparison of different parameter and parameter variance estimations: given are the average parameter (column 1) and parameter standard deviation (column 2) differences between 100 estimations of a four parameter transformation ($p_1$, $p_2$ for scale and rotation, $p_3$ and $p_4$ for translation). The difference in the computation is the stochastic model: one series was computed with the standard diagonal weight matrix $\mathbf{Q}_{hh}^{-1}$ neglecting the effects of high local correlations and the other series was computed using a weight matrix $\breve{\mathbf{Q}}_{hh}^{-1}$ where only every 2nd measurement was considered.

| Parameters | $(\breve{p}_N) - p_N \times 10^{-2}$ | $(\breve{\sigma}_{p_N}) - \sigma_{p_N} \times 10^{-2}$ |
|---|---|---|
| $p_1$ [-] | $0.020 \pm 0.545$ | $0.286 \pm 0.276$ |
| $p_2$ [-] | $-0.001 \pm 0.328$ | $0.250 \pm 0.251$ |
| $p_3$ [pixel] | $-0.043 \pm 8.529$ | $0.001 \pm 0.001$ |
| $p_4$ [pixel] | $0.149 \pm 7.026$ | $0.001 \pm 0.001$ |
| average | $0.031 \pm 2.767$ | $0.134 \pm 0.093$ |

Considering the effect of Gaussian smoothing, an approximation of an adjusted variance-covariance matrix $\mathbf{Q}_{hh}$ for a symmetric Gaussian distribution with $\sigma_G \geq 0.6$ is given by (*Berger*, 1999):

$$\breve{\mathbf{Q}}_{hh} = \frac{\sigma_0^2}{4\pi\sigma_G^2} \cdot e^{-\frac{\mathbf{D}_{xy}^2}{4\sigma_G^2}} \tag{3.21}$$

with $\mathbf{D}_{xy}^2$ being the squared pixel-to-pixel distance matrix. Thus for a Gaussian smoothing with $\sigma_G^2 = 1$, the maximum relative error in the diagonal elements of $\Delta\mathbf{P}$ becomes approximately $\pm 0.1$. With empirical tests one can show that for the parameter update $\delta\hat{\mathbf{p}}$ (as well as for the estimated parameters $\hat{\mathbf{p}}$ after all iterations) this transfers into an error on the order of $10^{-3}$ to $10^{-4}$ and for the parameter variance-covariance matrix $\mathbf{K}_{\hat{p}\hat{p}}$, this is on the order of $10^{-6}$ to $10^{-7}$. For the parameters, this error is at least one order of magnitude lower than what can be expected as the optimal accuracy of 0.05 pixel (*Ackermann*, 1984; *Zhang*, 2005). In case the adjusted patch scale is in the order of 1 (no scale change), the effects of the error in the weight matrix due to image interpolation is expected to be smaller (smaller correlation radius) than that of a Gaussian smoothing, and can thus be neglected.

The above estimates, however, only represent the effect of a pixel variance error due to interpolation and/or Gaussian smoothing. Although this is negligible, the effect of pixel-to-pixel correlation has not yet been addressed. The resulting parameter estimates and parameter standard deviations, when using $\breve{\mathbf{Q}}_{hh}$ (with correlated observations and random variance errors of $\pm 0.1$) instead of the standard diagonal $\mathbf{Q}_{hh}$ matrix, are summarized in Tabel 3.1: the influence on the estimated parameters as well as on the the respective parameter standard deviations is below the significance level.

Although the implementation of a corrected weight matrix is possible, a big drawback is the increased computation time. Because the $\mathbf{P}$-matrix is no more purely diagonal, the efficient row-wise computation of the normal matrix $\mathbf{N} = \mathbf{A}^T\mathbf{Q}_{hh}^{-1}\mathbf{A}$ (and $\mathbf{A}^T\mathbf{Q}_{hh}^{-1}\delta\mathbf{h}$) can no longer be carried out. In case Gaussian smoothing of $\sigma_G^2 = 1$ is applied, an adjacent

**Table 3.2:** Comparison of different parameter and parameter variance computations (for similarity transformations, 4 parameters) between a covariance matrix $\breve{\mathbf{Q}}_{hh}^{-1}$ with variable observation variances as well as pixel-to-pixel correlations (according to Eqn. (3.21) using $\sigma_G = 0.6$) and diagonal weight matrices $\mathbf{Q}_{hh_1}^{-1}$ to $\mathbf{Q}_{hh_3}^{-1}$. The three diagonal matrices differ in the number of pixels for which $\mathbf{Q}_{hh_1}(k,k) = 1$. For $\mathbf{Q}_{hh_1}^{-1}$, all pixels have equal weights, for $\mathbf{Q}_{hh_2}^{-1}$ two out of four neighboring observations have zero weight, and for $\mathbf{Q}_{hh_3}^{-1}$ three out of four neighbouring observations have zero weight. Shown are the differences with respect to the correlated weight matrix as well as the mean and the 1 $\sigma$ level of approximately 1800 realizations.

| Parameters $\times 10^{-2}$ | $\mathbf{Q}_{hh_1}$ | $\mathbf{Q}_{hh_2}$ | $\mathbf{Q}_{hh_3}$ |
|---|---|---|---|
| $p_1 - \breve{p}_1$ | $0.041 \pm 1.845$ | $0.010 \pm 2.101$ | $0.054 \pm 3.448$ |
| $p_2 - \breve{p}_2$ | $0.126 \pm 1.511$ | $0.136 \pm 1.689$ | $0.063 \pm 2.701$ |
| $p_3 - \breve{p}_3$ | $0.241 \pm 0.333$ | $0.237 \pm 0.345$ | $0.240 \pm 0.437$ |
| $p_4 - \breve{p}_4$ | $0.000 \pm 0.123$ | $0.000 \pm 0.134$ | $-0.004 \pm 0.203$ |
| average $\Delta p_N$ | $0.102 \pm 0.953$ | $0.096 \pm 1.067$ | $0.088 \pm 1.697$ |
| $\sigma_{p_1} - \breve{\sigma}_{p_1}$ | $0.001 \pm 0.000$ | $0.000 \pm 0.000$ | $0.000 \pm 0.000$ |
| $\sigma_{p_2} - \breve{\sigma}_{p_2}$ | $0.001 \pm 0.000$ | $0.001 \pm 0.000$ | $0.000 \pm 0.000$ |
| $\sigma_{p_3} - \breve{\sigma}_{p_3}$ | $0.257 \pm 0.147$ | $0.176 \pm 0.120$ | $0.007 \pm 0.161$ |
| $\sigma_{p_4} - \breve{\sigma}_{p_4}$ | $0.220 \pm 0.139$ | $0.152 \pm 0.109$ | $0.011 \pm 0.126$ |
| average $\Delta\sigma_{p_N}$ | $0.119 \pm 0.072$ | $0.082 \pm 0.057$ | $0.005 \pm 0.072$ |

pixel will have a correlation as high as 0.78. This is a strong smoothing for least-squares feature matching, but adding all the correlation effects, such a high pixel-to-pixel correlation in the local neighborhood might not be uncommon. To maintain the diagonal structure of the weight matrix without overestimating the parameter variances, a possible solution is to consider only every 2nd or even 3rd observation (pixel), compare Table 3.2. This reduces the maximum correlation to about $0.1 \cdot \sigma_0^2$ (with $\sigma_G^2 = 1$), while no loss in accuracy is expected, because only highly correlated observations are removed. Therefore, as long as the template and patch images for matching are big enough, the weight of every 2nd pixel will be assigned to zero.

The next Section describes the parameter estimation procedure based on Eqns. 3.12 to 3.16. It is to note here that in case of bad parameter convergence or large alterations (e.g., parameters with low significance due to insufficient image (observation) content), the parameters are introduced as stochastic variables in the estimation model itself. A priori information (e.g., from previous coarse matching results) thus helps to stabilize the parameter convergence and prevents parameters to drift into non-meaningful regimes.

### 3.2.2 Parameter Estimation

Eqns. 3.9, 3.12, and 3.16 define the affine parameter estimation scheme. In principle, however, the set of parameters can be reduce to two (translation components) or increased to eight (projective transformation). The set of parameters that are estimated depends on the problem type, i.e., there should be enough parameters to comprehensively model the distortions, while no over-parametrization should result. According to *Gruen* (1985), non-determinable parameters have an impairing effect on the estimation process and worsen the matching quality. It is thus important to perform a parameter determinability test during the iterations of the non-linear system. The parameter influence can for example be estimated by computing its relative contribution to the cofactor matrix $\mathbf{Q}_{\hat{p}\hat{p}}$. With $q_{ij}$ being the element $(i, j)$ of $\mathbf{Q}_{\hat{p}\hat{p}}$, its contribution is given by (*Koch*, 1988):

$$\delta_{ii} = \frac{\sum_j q_{ij}^2}{q_{ii} \cdot \sum_j q_{jj}^2} \tag{3.22}$$

In case $\delta_{ii}$ is high, the parameter $p_i$ strongly correlates with at least one other parameter and thus should be excluded or combined (e.g., constrained) with an other parameter. Another possibility to check for correlated parameters is to compute the eigenvalues of the normal matrix $\mathbf{N} = \mathbf{A}^T\mathbf{PA}$. By solving the eigenvalue problem, the independent components of the parameter space can be used to exclude those, whose influence is negligible (i.e., having a small eigenvalue). To constrain a parameter $p_i$, the inversion of the normal matrix is computed by a constraint Moore-Penrose inverse, $\mathbf{N}^{-1} \to \mathbf{N}^{+}$ (*Moore*, 1920), with matrix $\mathbf{C}$ containing the constraints of the corresponding parameters:

$$\begin{pmatrix} \mathbf{N}^{+} & \cdots \\ \vdots & \ddots \end{pmatrix} = \begin{pmatrix} \mathbf{N} & \mathbf{U} & \mathbf{C}^T \\ \mathbf{U}^T & 0 & 0 \\ \mathbf{C} & 0 & 0 \end{pmatrix}^{-1} \tag{3.23}$$

where matrix $\mathbf{U}$ contains all eigenvectors of $\mathbf{N}$, whose eigenvalues are below a given threshold. Depending on the image content, parameters have to be constrained a priori in case they cannot be determined.

### 3.2.3 Outlier Detection

The detection of outliers or blunders is a critical task in the process of LSM. Erroneous observations will cause unexpected outcomes and incorrect results (e.g. *Zhllin et al.*, 2001). The sensitivity of parameter estimation with respect to discrepancies between the template and the patch image can be detected using the concepts of internal, external reliability, as well as data snooping (*Baltsavias*, 1991). The problem, however, is not trivial because any gross error influences all the residuals in the least-squares adjustment process. Thus the methods presented below rely on the principle of very few outliers (compared to the total number of observations). If this prerequisite is not given, inconsistencies due to a wrong model assumption, an erroneous cofactor matrix of the observations, or gross errors can hardly be detected, if at all.

Gross error detection is based on the principles of null-hypothesis testing that is performed after a first solution has been found. The residuals $\hat{\mathbf{v}}$ in Eqn.(3.5) are then computed as:

$$\hat{\mathbf{v}} = \tilde{\mathbf{g}}(\hat{\mathbf{p}}, \mathbf{g}) - \mathbf{h} = \mathbf{A} \cdot \delta\hat{\mathbf{p}} - \delta\mathbf{h} \tag{3.24}$$

$$\text{whereas} \quad E\langle\hat{\mathbf{v}}\rangle = 0, \quad \text{and} \quad E\langle\hat{\mathbf{v}}\hat{\mathbf{v}}^T\rangle = \sigma_0^2 \cdot \mathbf{Q}_{hh}^{-1} \tag{3.25}$$

To test the assumed normal distribution (global test) of $\hat{\mathbf{v}}$, the null-hypothesis $H_0$ is set up that $E\langle\hat{\sigma}_0^2\rangle = \sigma_0^2$ is true, if the test criterion $\Theta$ is distributed as the central Fisher distribution:

$$\Theta = \frac{\hat{\sigma}_0^2}{\sigma_0^2} \tag{3.26}$$

If $\Theta$ exceeds the test criterion $F(1-\alpha, r, \infty)$ (Fisher distribution with degree of freedom $r$ and significance level $\alpha$), $H_0$ is rejected. To compute $\sigma_0^2$, an a priori estimate of the noise content with respect to the involved images (compare Eqn. 3.4) has to be known. Because generally there is no information about this quantity a priori, the goodness of the fit cannot reliably be expressed with this global test. As proposed in *Berger* (1999), the normalized cross correlation in the adjusted state is used as a quality indicator:

$$\rho = \frac{\sum(\mathbf{h} - \bar{\mathbf{h}}) \cdot \sum(\tilde{\mathbf{g}} - \bar{\tilde{\mathbf{g}}})}{\sqrt{\sum(\mathbf{h} - \bar{\mathbf{h}})^2 \cdot \sum(\tilde{\mathbf{g}} - \bar{\tilde{\mathbf{g}}})^2}} \tag{3.27}$$

where $\bar{\mathbf{h}}, \bar{\tilde{\mathbf{g}}}$ indicate the mean value of the template and the adjusted patch image, respectively. If the matching is perfect, $\rho$ is 1. A value clearly below 1 either indicates a wrong model assumption, an incorrect weight matrix, and/or erroneous observations. Assuming that the model complexity is sufficient for the deformation scenarios and that the initial estimates (see Section 3.2.5) were well chosen, the most probable issue causing the test to fail are errors in the observation vector (outliers). Therefore, additional tests for gross error detection must be applied. An estimate of the internal reliability is obtained by computing the smallest detectable gross error with probability $\beta(\lambda) = 1 - P(\lambda)$, $P(\lambda)$ being the probability of accepting a false hypothesis, $\lambda$ being the non-centrality parameter. For a diagonal weight matrix $\mathbf{P}$, and subscript $i$ as its $i$-th element, this is (*Baarda*, 1967; *Gruen*, 1978):

$$\lambda_1^i = \sigma_0 \cdot \Big(\frac{\lambda_1}{\mathbf{P}_i^2 q_{v_i v_i}}\Big)^{1/2} \tag{3.28}$$

with $q_{v_i v_i}$ being the i-th diagonal element of $\mathbf{Q}_{vv} = \mathbf{Q}_{hh} - \mathbf{A}\mathbf{Q}_{\hat{p}\hat{p}}\mathbf{A}^T$. Due to the high redundancy, the internal reliability typically is very high.
Also if $H_o$ is rejected, a test on the residuals can be used to detect outliers (typically called data snooping, see *Gruen* (1978)). The test variables are the standardized residuals (*Baarda*, 1967; *Carosio*, 2008):

$$w_i = \frac{\hat{v}_i}{\hat{\sigma}_{v_i}} = \frac{\hat{v}_i}{\hat{\sigma}_0 \cdot \sqrt{q_{v_i v_i}}} \tag{3.29}$$

The acceptance interval for $w_i$, with the type I error size $\alpha$ being typically chosen as $\alpha = 0.05$ or $\alpha = 0.01$, is given by:

$$-F^{1/2}(1-\alpha, 1, r) < w_i < F^{1/2}(1-\alpha, 1, r) \tag{3.30}$$

with $r$ being the system redundancy. As shown by *Carosio* (2008) or *Koch* (2012), this is the same as the t-distribution defined as:

$$|w_i| < t(1-\alpha/2, r) \tag{3.31}$$

For $\alpha = 0.05$ and $\alpha = 0.01$, $|w_i|$ yields a critical value of 1.960 and 2.576, respectively (for $r \to \infty$).

Practically, several outliers can be detected simultaneously after a first convergence of the iteration process, if there are multiple large residuals $w_i$ that are geometrically independent (*Luhmann et al.*, 2006). If the global normalized cross correlation test indicates possible gross errors, all standardized residuals are validated with respect to their critical test value (Eqn. (3.31)). After detecting these outliers, the iterative matching process is restarted with a modified weight matrix, i.e., the weighting coefficients of the erroneous observations are set to zero. For each iteration run, the maximum number of observations assigned as outliers is restricted and depends on the total number of observations available. This iterative procedure is repeated until the global correlation test indicates no further errors. It is to note that after each parameter convergence, all observations are again tested for being gross errors, whereas for each cycle, the maximum number of assigned outliers is increased. Following this procedure allows false positive outliers to be again considered as valid observations. As shown in Section 3.2.6, this procedure makes the traditional least-squares image matching a relatively robust parameter estimation procedure, even though it is a sensitive $L_2$−Norm minimization.

Because the components of matrix $\mathbf{A}$ are computed using a 2-dimensional convolution kernel of length 5 (Section 3.2.1), the image gradient approximations of neighboring pixels are strongly correlated. In case of a detected outlier, the weight elements of its neighboring pixels are also set to zero (as the computed gradient components are strongly influenced by the erroneous observation). This procedure was found to have a stabilizing effect for the parameter estimation process.

### 3.2.4 Estimation Scheme

Although the algorithm explained in Sections 3.2.1 to 3.2.3 is fully automatic, a couple of choices involved in the principle of the LSM process have to be made and adjusted depending on the problem type. For monitoring surface motion in permafrost areas, the displacement of individual rock boulders is of interest and LSM is carried out for predefined feature points (Section 3.3.1). For this configuration, the criteria in the following two paragraphs have to be addressed.

#### Search Radius and Matching Window

The first task for feature matching between two images is to define a search radius, i.e., a maximum expected difference between the image coordinates of the template and patch images, respectively. In principle the complete image can be used to search for the best match. In practical applications, however, the displacement to be expected is not random and a rough estimate is usually known a priori. This upper maximum displacement limit can be used to define a reduced search radius, thus speeding up the matching process. As soon as the initial estimate of the transformation has been obtained and applied (Section 3.2.5), the actual size of the matching window has to determined. *Debella-Gilo and Kääb* (2012) made an attempt to adaptively adjust the window size for the normalized cross-correlation procedure. They used image signal-to-noise ratio estimates to find suitable matching candidates and correlation peaks as a function of the template window size for the optimal dimension of the matching window. For the LSM used here, the dimensions of the matching windows are defined individually by evaluating the determinability of the parameters before the iteration starts. With respect to the selected mathematical model, the window dimensions (i.e., the number of observations) are tested using the eigenvalues of the normal matrix $\mathbf{N} = \mathbf{A}^T\mathbf{PA}$. Solving the eigenvalue problem projects the normal matrix components into an orthogonal system with eigenvalues characterizing the individual contributions. To reliably solve for the parameters, the difference between the largest and smallest eigenvalue should not exceed a certain limit. Because this is only an approximation (e.g. possible outliers are considered as valid observations and only the initial transformation is known), the limit should be set rather optimistically.

Because a rough topography (like in case of rock glaciers and similar permafrost creep areas) provides ideal targets for feature tracking (high local contrast), the window sizes determined hereby are typically in the order of 10 by 10 pixel (for ideal targets). When matching is performed between epochs with different lighting conditions (e.g. between images taken in the evening and morning), there is the danger of estimating motion due to the varying reflecting surface rather than true motion (see Fig. 3.2). Thus the LSM window size is the maximum of either a user-defined minimum window dimension or given by the minimum dimension that is required to reliably determine all parameters.

The minimum matching window thus has to be a compromise between the estimation of local deformations and the robustness to variable illuminations. The optimal window size with regard to both aspects is hard, if not impossible to determine. For the purpose of this study, a minimum window size of $30 \times 30$ pixels was used.

#### Stepwise Simplification of Model Complexity

An initial choice of the model complexity (projective, affine, similarity, or only translation components), suitable for accurately estimate the expected transformation, has to be made. For steady scene flows or images taken from very large distances with respect to the feature size (e.g., satellite images), the simplest model (translation components) is usually a good choice and comparable to the widely used normalized cross-correlation

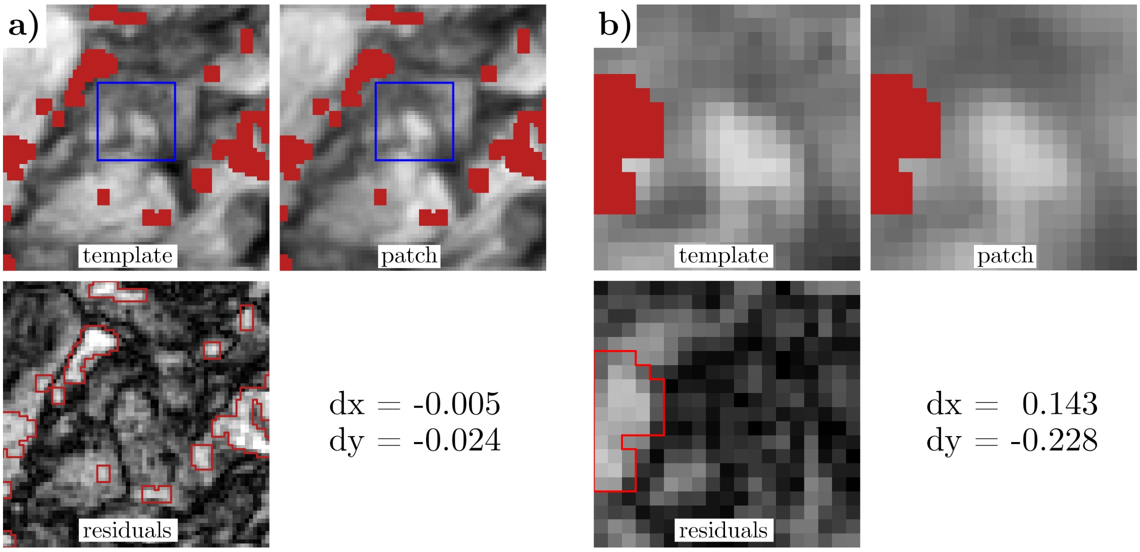


**Figure 3.2:** Difference in estimated translation components with respect to the size of the matching window. The template and the patch images were acquired with $\delta t = 10h$. Although only diffuse light scattering illuminated the scenes, some structures in the image changed in appearance. (a) Shows the template, patch, and residual images for a large matching window ($60 \times 60$ pixels), whereas in (b) a small matching window ($19 \times 19$ pixels) was used. The true motion between these epochs is in the order of $10^{-2}$ pixel. The respective translation components (dx and dy) show that the effect of scene illumination for small window dimensions easily exceeds the precision of the LSM by up to one order of magnitude. Red areas are gross errors detected during the adjustment (residual images show the respective contours). The residual images show the absolute differences between the template and the adjusted patch image, amplified by a factor 2. The blue square in (a) illustrates the matching window used in part (b).

motion estimation (e.g. *Kääb and Vollmer*, 2000) but with intrinsic sub-pixel accuracy. For objects exposed to simple rotation and/or scaling, a 4-parameter similarity transformation is applicable. For images taken in the vicinity of the moving object, the estimation of additional transformation parameters can be helpful, specially if rock or ice boulders are expected to rotate (in three dimensions). For stereo/multi-view matching (Section 4.3) or other projection scenarios, an 8-parameter transformation matrix is suitable.

Another argument for this approach can be made in scenarios with notable effects of geometric projection. In case there is a flat angle between the line of sight and the direction of motion (or if the direction of motion changes significantly between two measurement epochs), matching the two projected images is only well modelled if a projective or at least an affine transformation is assumed. Thus the matching window along with the image content have to allow the estimation of these parameters. On the other hand, if the flow field is expected to be inhomogeneous, the matching window should be as small as possible. This is also in good aggrement with the fact that approximations of uniform displacements for larger window dimensions is violated in many situations (e.g. *Whillans and Tseng* (1995); *Debella-Gilo and Kääb* (2012)).

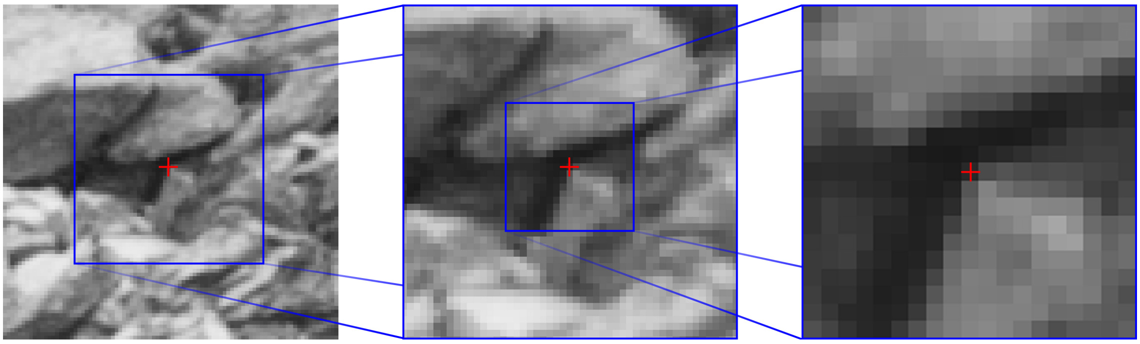

**Figure 3.3:** Principle of stepwise model simplification and estimation of local displacements. For example the most complex transformation model is estimated using a large window (left image), whereas the simplest model with a small window is estimated in the last step, consequently better describing the motion of the feature point itself. The red cross indicates the selected feature point for which the transformation is estimated.

As noted in the previous paragraph, the minimum matching window dimension is defined either by the minimum dimensions allowed or by the image content. As more parameters require more observations to be included, this naturally leads to larger window dimensions. For high accuracy matching in complex flow regimes (e.g., the study of interannual variations in rock glacier kinematics (e.g. *Delaloye et al.*, 2008a) and spatially variable flow characteristics (e.g. *Delaloye et al.*, 2008b)), however, a matching window that is as small as possible is favorable. Thus a step-wise estimation of parameters from high to low model complexity can be applied (Fig. 3.3). The adaptivity in the process is realized by iteratively reducing the model complexity from the given starting complexity (usually affine) after the current parameter estimation has finished. By reducing the number of parameters to be estimated, smaller matching windows may be possible and thus the eigenvalues are again evaluated for the new configuration. A matching procedure can be started, for example, by estimating an affine transformation, followed by a similarity transformation (with a smaller matching window), and a two-parameter translation transformation, again with a smaller matching window. Mathematically this reads as:

$$\begin{pmatrix} x \\ y \\ 1 \end{pmatrix}_k^T = \begin{pmatrix} u \\ v \\ 1 \end{pmatrix}_k^T \cdot \underbrace{\begin{pmatrix} p_1^a & p_2^a & 0 \\ p_3^a & p_4^a & 0 \\ p_5^a & p_6^a & 1 \end{pmatrix}}_{\text{affine}} \cdot \underbrace{\begin{pmatrix} p_1^b & -p_2^b & 0 \\ p_2^b & p_1^b & 0 \\ p_5^b & p_6^b & 1 \end{pmatrix}}_{\text{similarity}} \cdot \underbrace{\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ p_5^c & p_6^c & 1 \end{pmatrix}}_{\text{translation}} \tag{3.32}$$

with superscripts $a, b$, and $c$ indicating the different mathematical models. If this estimation procedure is applied, outlier detection (Section 3.2.3) is carried out only in the first (most complex) model with high redundancy. It was observed that for well sampled images this stepwise decrease of model complexity does not yield a significant improvement. This indicates a coherent transformation for the dimension of the LSM window used in the high order model. In more difficult scenarios of lower image contrast and complex local deformations, the proposed matching strategy might be of advantage. During the model simplification, the estimated parameters from the previous model are used as initial guess, whereas those not being estimated anymore are applied in a pre-multiplication. If, for example, an affine transformation was estimated in a first step, the set of parameters for the similarity transformation $(p_1^b, p_2^b, p_5^b, p_6^b)$ will be estimated as:

$$\begin{aligned}
\begin{pmatrix} x \\ y \\ 1 \end{pmatrix}_k^T &= \begin{pmatrix} u \\ v \\ 1 \end{pmatrix}_k^T \cdot \mathbf{T}^a \cdot \begin{pmatrix} p_1^b & -p_2^b & 0 \\ p_2^b & p_1^b & 0 \\ p_5^b & p_6^b & 1 \end{pmatrix} \\
\text{where} \quad \mathbf{T}^a &= \begin{pmatrix} p_1^a & p_2^a & 0 \\ p_3^a & p_4^a & 0 \\ 0 & 0 & 1 \end{pmatrix}
\end{aligned} \tag{3.33}$$

The initial guess for the parameters $p_1^b$ and $p_2^a$ is 1 and 0, respectively, whereas for $p_5^b$ and $p_6^b$, the respective translation components from the previous estimation, $p_5^a$ and $p_6^a$, are used. The simplest model then is estimated as:

$$\begin{aligned}
\begin{pmatrix} x \\ y \\ 1 \end{pmatrix}_k^T &= \begin{pmatrix} u \\ v \\ 1 \end{pmatrix}_k^T \cdot \mathbf{T}^{a,b} \cdot \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ p_5^c & p_6^c & 1 \end{pmatrix} \\
\text{where} \quad \mathbf{T}^{a,b} &= \begin{pmatrix} p_1^a & p_2^a & 0 \\ p_3^a & p_4^a & 0 \\ 0 & 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} p_1^b & -p_2^b & 0 \\ p_2^b & p_1^b & 0 \\ 0 & 0 & 1 \end{pmatrix}
\end{aligned} \tag{3.34}$$

In this case, the final transformation matrix is given by:

$$\begin{aligned}
\begin{pmatrix} x \\ y \\ 1 \end{pmatrix}_k^T &= \begin{pmatrix} u \\ v \\ 1 \end{pmatrix}_k^T \cdot \mathbf{T}^{a,b,c} \\
\text{with} \quad \mathbf{T}^{a,b,c} &= \begin{pmatrix} \hat{p}_1 & \hat{p}_2 & 0 \\ \hat{p}_3 & \hat{p}_4 & 0 \\ \hat{p}_5 & \hat{p}_6 & 1 \end{pmatrix} = \begin{pmatrix} p_1^a p_1^b + p_2^a p_2^b & -p_1^a p_2^b + p_2^a p_1^b & 0 \\ p_3^a p_1^b + p_4^a p_2^b & -p_3^a p_2^b + p_4^a p_1^b & 0 \\ p_5^c & p_6^c & 1 \end{pmatrix}
\end{aligned} \tag{3.35}$$

Regarding the parameter variance estimation, error propagation must be applied for the combined parameter sets:

$$\mathbf{K}_{\hat{p}\hat{p}} = \begin{pmatrix} \mathbf{K}_{pp}^{a,b} & \mathbf{0} \\ \mathbf{0} & \mathbf{K}_{pp}^{c} \end{pmatrix} \tag{3.36}$$

$$\text{with} \quad \mathbf{K}_{pp}^{a,b} = \mathbf{F}^T \cdot \begin{pmatrix} \mathbf{K}_{pp}^{a} & \mathbf{0} \\ \mathbf{0} & \mathbf{K}_{pp}^{b} \end{pmatrix} \cdot \mathbf{F} \tag{3.37}$$

$$\text{and} \quad \mathbf{F}^T = \begin{pmatrix} p_1^b & 0 & 0 & p_2^b & p_1^a & p_2^a \\ -p_2^b & 0 & 0 & p_1^b & p_2^a & -p_1^a \\ 0 & p_2^b & p_1^b & 0 & p_3^a & p_4^a \\ 0 & p_1^b & -p_2^b & 0 & p_4^a & -p_3^a \end{pmatrix} \tag{3.38}$$

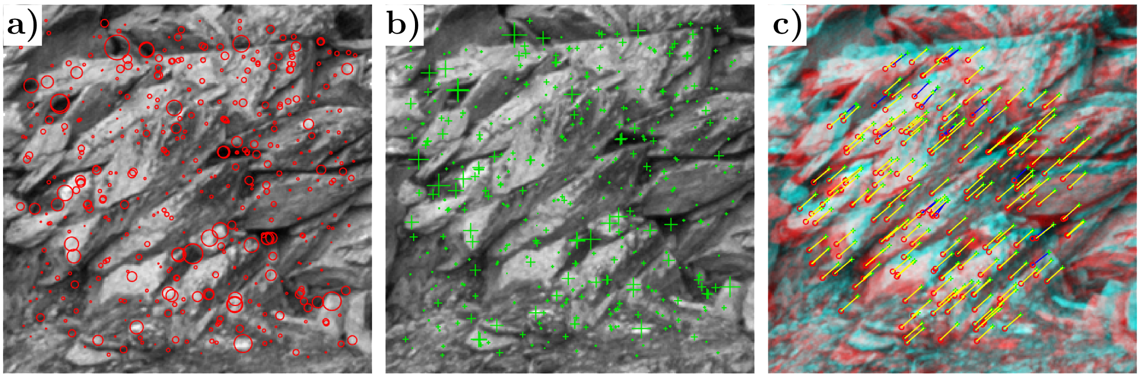


**Figure 3.4:** (a), (b) Detected SURF features in the template and patch image, respectively. The size of the markers indicate the strength of the feature points. (c) Matched features between the template and the patch image. Yellow lines indicate all matching pairs within the 3 pixel tolerance found after RANSAC transformation estimation, blue lines indicate false matching pairs. To approximate an optimal transformation around the point of interest, only SURF features within a certain radius are considered for the estimation.

### 3.2.5 Initial Estimates

As the parameter estimation of non-linear systems needs to be carried out iteratively using a linearized formulation of the mathematical model (Section 3.2.1), it is crucial to have good initial parameter estimates. For the least-squares matching, *good* means that the initial parameters have to be in the pull-in range of about 3 pixel, depending on the image content (*Baltsavias*, 1991). The following paragraphs describe the implemented initial parameter estimation routines: (1) using surface descriptors coupled with the RANSAC estimation procedure (*Fischler and Bolles*, 1981), (2) hierarchical matching using image pyramids, and (3) Normalized Cross-Correlation matching. For automated matching, these three methods are carried out with priority as listed. In case one fails or the estimated fit is not satisfying, the next method is tested.

#### Surface Descriptor Matching

Surface point descriptors are used to describe image features in two images. A robust matching is applied to find a common set of points between these images (i.e., a common transformation is estimated to match one set of points with the others). Here, SURF descriptors are used (*Bay et al.*, 2008), a fast algorithm to reliably detect and describe scale and rotation invariant interest points. The descriptor is based on the sum of first-order Haar wavelet responses within the feature sub-regions, whereas the typical length of the descriptor vector is 64. For matching, the trace of the Hessian matrix (sign of Laplacian) is used to speed up the process, as matching is tested only between features of equal sign. SURF descriptors are similar to SIFT descriptors (*Lowe*, 2004) but faster and more robust.

Figure 3.4, a) and b), show SURF features detected in the template and patch image, respectively. Once these descriptors for both images have been determined, a descriptor matching is performed for plausible point correspondences: for every descriptor in the patch image, the most probable corresponding descriptors in the template image are computed. The interest points are then tested using the RANSAC estimation procedure with a point matching tolerance of 2 pixels (to allow for some inaccuracies and local deformations). Depending on the number of successful matches, the valid radius around the point of interest is reduced to get a more accurate approximation of the local transformation. Specially in situations, where deformations show a high spatial variability, the latter restriction along with a higher tolerance value often lead to a successful initial estimate.

### Hierarchical Matching

The principle of hierarchical image matching is relatively simple and extensively described for example in *Baltsavias* (1991) or *Zhang* (2005). The idea is to scale down the image around the point of interest such that the maximum expected motion can be estimated using the iterative adjustment with initial guess parameters indicating no transformation. After the first parameter convergence, the next scale is used for matching, whereas the initial parameters are scaled according to the scale difference (in fact, only the translation components need to be scaled, as the other parameters are scale invariant). This process is repeated until the original resolution has been reached.

### Normalized Cross-Correlation

Due to its simplicity and fast computation, normalized cross-correlation (and some improved versions of that principle) is a very popular image matching technique also used to directly derive motion vectors of glaciers and other creeping landforms (*Kääb and Vollmer*, 2000; *Heid and Kääb*, 2012). For each relative shift, $i, j$ between the patch image $\mathbf{g}$ and the template image $\mathbf{h}$ a correlation value is computed as:

$$\mathbf{NCC}(i,j) = \frac{\sum_{x,y}\left(\mathbf{g}(i+x,j+y)-\bar{\mathbf{g}}\right)\cdot\left(\mathbf{h}(x,y)-\bar{\mathbf{h}}\right)}{\sqrt{\sum_{x,y}\left(\mathbf{g}(i+x,j+y)-\bar{\mathbf{g}}\right)^2\cdot\sum_{x,y}\left(\mathbf{h}(x,y)-\bar{\mathbf{h}}\right)^2}} \tag{3.39}$$

where $x, y$ is the position in the template image and $\bar{\mathbf{h}}$ and $\bar{\mathbf{g}}$ are the mean intensity values of the template and patch image, respectively. The best match at offset $i, j$ is given by the highest correlation value (see Fig. 3.5). An alternative fast computation of this spatial filter is to compute the cross-correlation matrix in the frequency domain, i.e., multiplying the Fourier transform of the template image with the conjugated Fourier transform of the patch image (e.g. *Lewis*, 1995; *James H. McClellan*, 2003). Because normalization of luminosity is lost hereby, a better solution is to only consider the phase information (as differences in image intensities are contained in the amplitude part), typically called phase correlation. A full review on existing image matching techniques based on the cross-correlation principle can be found in *Heid and Kääb* (2012).

The principle of cross-correlation can also be extended to subpixel accuracy with additional parameters for higher order transformations. *Zheltov and Sibiryakov* (1997) showed that this results in an equivalent formulation as obtained by the LSM technique.

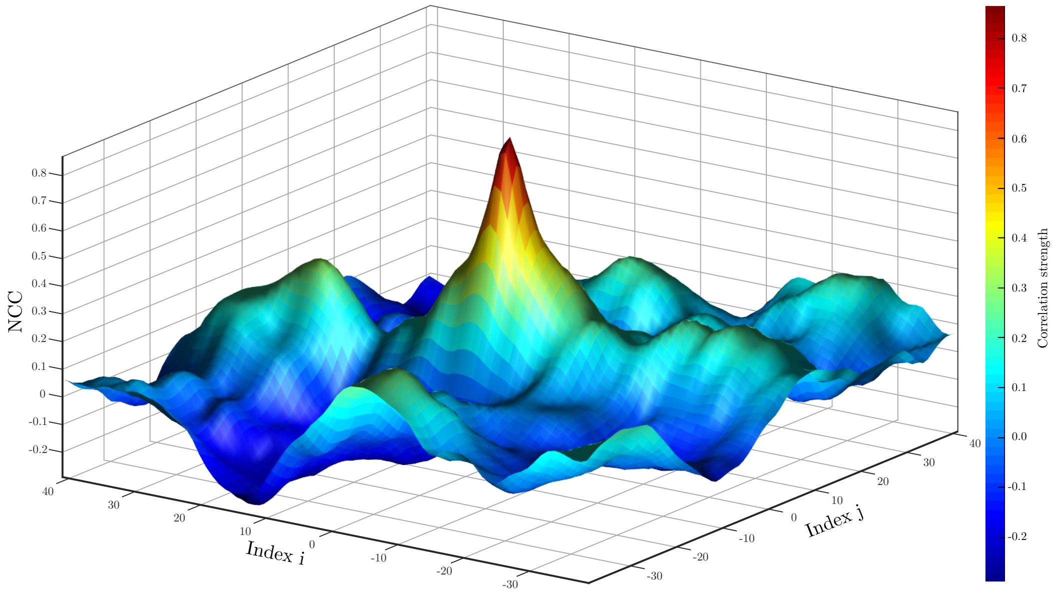


**Figure 3.5:** Surface plot of normalized cross correlation estimates (Eqn. 3.39) with the offsets $i$ and $j$ ranging from $-40$ to $+40$. The initial guess for the translation components is found at the offset, where $NCC$ is maximal.

### 3.2.6 Robust Estimation

The least-squares adjustment introduced in Section 3.2.1 is an optimal estimator (maximum likelihood estimation, MLE) in case the stochastic behavior of the observations is based on an unbiased symmetric distribution of residuals. For this to be true, no gross errors and no systematic errors are allowed to be present and the residuals after the adjustment must be randomly distributed. The MLE is a non-robust estimator as it is very sensitive to gross errors (i.e., a single gross error can have a large impact on the parameters being estimated such that time consuming outlier detection and elimination procedures have to be applied, Section 3.2.3). A robust estimator has to be robust to non-Gaussian data, to gross errors in the data, and it has to give the optimal solution while still being efficient. There are different robust estimators, typically described by the influence function and/or the breakdown point (e.g. *Carosio*, 2001). The latter being the more popular indicator, it can be shown that the maximum breakdown point (defined as the maximum portion of outliers that can be tolerated without significantly influencing the estimator) is equal to 50%, which is known as the $L_1 - norm$, or median.

Many types of robust estimators exist, whereof the M-estimator class (maximum likelihood estimators) is among the most popular ones. Without going into detail, the principle of these estimators is to define a loss function $\rho(\hat{v})$, with $\rho$ defining, how a specific residual and its corresponding observation will be treated in the next iteration. Practically the estimation is carried out by performing an initial traditional Least-Squares adjustment. Based on its residuals, the estimation is repeated with re-weighted observations, i.e. for

Huber's method (*Huber*, 1964):

$$\sum_{i=1}^{n} \mathbf{p}_{ii} \cdot \rho(\hat{v}) \rightarrow min \tag{3.40}$$

$$\text{with} \quad \rho(\hat{v}) = \begin{cases} \frac{1}{2}\hat{v}^2 & \text{for} \|\hat{v}\| < c \\ c \cdot \|\hat{v}\| - \frac{1}{2}c^2 & \text{for} \|\hat{v}\| \geq c \end{cases} \tag{3.41}$$

$$\text{and} \quad c = const. \tag{3.42}$$

This leads to an iterative optimization according to:

$$\delta\hat{\mathbf{x}}^0 = \left(\mathbf{A}^T \cdot \mathbf{Q}_{ll}^{-1} \cdot \mathbf{A}\right)^{-1} \cdot \mathbf{A}^T \cdot \mathbf{Q}_{ll}^{-1} \cdot \delta\mathbf{l} \tag{3.43}$$

$$\hat{\mathbf{v}}^k = \mathbf{A} \cdot \delta\hat{\mathbf{x}}^k - \delta\mathbf{l} \quad k \text{ being 0 in the first iteration} \tag{3.44}$$

For the next iterations, a re-weighting matrix is defined:

$$\mathbf{W}^k = \operatorname{diag}\left(\rho(\hat{v}_1^k) \quad \rho(\hat{v}_2^k) \quad \dots \quad \rho(\hat{v}_n^k)\right) \tag{3.45}$$

$$\delta\hat{\mathbf{x}}^k = \left(\mathbf{A}^T \cdot \mathbf{W}^k \cdot \mathbf{Q}_{ll}^{-1} \cdot \mathbf{A}\right)^{-1} \cdot \mathbf{A}^T \cdot \mathbf{W}^k \cdot \mathbf{Q}_{ll}^{-1} \cdot \delta\mathbf{l} \tag{3.46}$$

The procedure is repeated until $\delta\hat{\mathbf{x}}$ has converged. For the LSM being a non-linear adjustment in general, this principle implies a second (sub-) iteration loop. Following *Dumouchel and O'Brien* (1991), a robust adjustment was implemented and tested against the traditional LSM technique with iterative gross error elimination.

### 3.2.7 Performance

Table 3.3 gives an overview of a set of performance tests that were conducted using different synthetic transformations with six parameters (affine transformation). The transformations were applied to different image windows, randomly positioned in the image. For good performance, the difference between the applied transformation and the re-estimated transformation should be minimal. To reliably estimate the average quality, the test was repeated several thousand times. In total, three test were performed:

(A) Zero noise: simple re-estimation of synthetic transformations.

(B) With noise: both images were contaminated with different noise (i.e., White noise of random local variance for the patch image and Poisson noise for the template image, respectively, were used).

(C) With noise and gross errors: the same noise contamination as in (B) but with additional randomly distributed patches of varying pixel intensities. The error patches were generated in a random way, of both multiplicative and additive nature.

For all these tests, both, the LSM technique with iterative gross error detection and elimination, as well as the robust matching strategy using the Huber loss function defined in Eqn. (3.41) with $c = 1.345$ were applied. As there were no systematic errors involved, the mean value of the parameter differences were all found to be in the range of $10^{-3} - 10^{-5}$ and are thus not listed in Table 3.3. Both methods yield very good results for case (A) and case (B), with a $1\sigma$ level below 0.05 pixel for the translation components (note that

**Table 3.3:** Synthetic test of feature matching performance using the Ordinary Least-Squares (OLS) and the Robust Least-Squares (RLS) procedure. The parameters correspond to an affine transformation defined in Eqn. (3.14). A match was considered successful, if a quality value of more than 0.8 was reached (Eqn. 3.27). For more explanations, see text.

| | case A | | case B | | case C | |
|---|---|---|---|---|---|---|
| | OLS | RLS | OLS | RLS | OLS | RLS |
| $\sigma(p_{1,syn} - p_{1,est})$ | 0.0007 | 0.0005 | 0.0038 | 0.0031 | 0.0064 | 0.0063 |
| $\sigma(p_{2,syn} - p_{2,est})$ | 0.0004 | 0.0003 | 0.0027 | 0.0023 | 0.0050 | 0.0049 |
| $\sigma(p_{3,syn} - p_{3,est})$ | 0.0004 | 0.0003 | 0.0030 | 0.0029 | 0.0054 | 0.0054 |
| $\sigma(p_{4,syn} - p_{4,est})$ | 0.0006 | 0.0004 | 0.0034 | 0.0032 | 0.0058 | 0.0059 |
| $\sigma(p_{5,syn} - p_{5,est})$ | 0.0070 | 0.0061 | 0.0350 | 0.0344 | 0.0462 | 0.0471 |
| $\sigma(p_{6,syn} - p_{6,est})$ | 0.0089 | 0.0075 | 0.0362 | 0.0345 | 0.0860 | 0.0797 |
| # valid matches | 2464 | 2443 | 2360 | 2362 | 1286 | 1428 |

the other parameters are scale invariant thus showing a different error level). Due to the necessary image interpolation and transformation step, small errors are also observed for the noise-free case (A). For images with variable outliers, case (C), both techniques give results at the expected 0.05 pixel level (*Danuser*, 1996), with the robust least-squares showing a higher number of positive matches. Considering this result, both methods are assumed to be good estimators for the matching techniques, where the robust method is preferred as the number of positive matches is slightly higher.

## 3.3 Data Processing Strategy

There are different possible scenarios of how to process image sequences in order to optimally extract a continuous displacement estimate over time. Generally, the image sampling rate should be well above the minimum time resolution needed to capture the processes of interest. Further it is to note that the time separation between two matching epochs should not be larger than the expected period, where reliable detection can be performed (i.e., it depends on the expected possible tracking period and surface changes). For the study of the Grabengufer permafrost creep (see Chapter 2), a sampling interval of one image per day was found to be sufficient, as the projected movement is in the order of 0.02 pixel per day (see Appendix E).

In principle, every image can be defined as a template for the forthcoming images to estimate feature trajectories in the image plane. When processing 100 images for example, there is a total of $(100 - 1) \cdot 100/2 = 4950$ independent image pairs. For the purpose of this work, only high-quality images with a minimal snow coverage (see Section 3.3.3) are considered for a template candidate. A time separation of roughly 10-30 days for a new template image is used, whereas feature tracks are being estimated for a time span of up to two years after the template epoch.

### 3.3.1 Feature Point Selection

Image matching is optimally performed in areas of good contrast with image gradients in different directions (e.g. *Baltsavias*, 1991). Feature points for tracking are selected using a corner point detector with minimum contrast and quality threshold (*Harris and Stephens*, 1988). For matching and reconstruction of a scene, a homogeneous distribution is preferred, as the matching points will be available in all image regions. For this to become true, the image is segmented into equal areas, whereas the number of points in each area is adjusted accordingly.

### 3.3.2 Image Preprocessing

For monitoring ice- or rock glaciers as well as landslides in remote areas, cameras are usually installed such that images are captured at constant time intervals during the day (e.g. *Travelletti et al.*, 2012; *Bernard et al.*, 2013). Even if there is the possibility to control the camera remotely, images are usually also captured in situations where the scenery is dominated by fog, snowfall, or when ice crystals cling to the lens (e.g. *Rüfenacht et al.*, 2014). If there is no adaptive control for deciding whether or not to capture an image, each picture must first be analyzed for its valuable data content. Therefore the first step is to sequentially evaluate the content of the image for its quality. The current procedure applies hierarchical tests that each image has to pass in order to be considered valid for further processing. The test series includes:

A) Capture time window selection: for the current motion speed of the permafrost creep, one image per day is enough to capture even the fastest processes in view, thus images taken in the morning hours are preferred as the lightening conditions turned out to be relatively stable. This minimizes the varying reflectance effect (though there are still large variations present, see Section 3.1.2), and avoids the problem of shadow casting, which produces strong local luminosity gradients that potentially degrade the matching quality.

B) Image contrast check: images are tested for contrast in order to be considered valid. A lack thereof is usually an indication of no visibility due to fog or bad weather, water droplets or accumulated ice crystals on the lens cover glass. For the two cameras in use (see Chapter 6), the global threshold for low contrast image content was set such that the luminosity values must span a minimum of half the dynamic range (e.g., 128 levels for an image recorded with 8 bit resolution). Images that were taken with an exposure either set too high or too low are rejected, as the measured image mean deviates considerably from its expected value. Also the image standard deviation is used to roughly decide, if there is enough image contrast in the global scale.

C) Image noise check: here a simple method to estimate image noise is used for evaluation. Images, with a high standard deviation of the upper part of the histogram, are be rejected. The upper part of the histogram is computed by subtracting a median filtered image from its original input.

D) Snow cover check: snow cover is estimated for each image (see Section 3.3.3), i.e., areas covered by snow are detected and, if the area is mostly covered by snow, the image will be rejected.

Images passing these evaluations are equalized in terms of their histograms. Contrast and illumination are analyzed and adjusted for each image being further processed such that equal objects appear in approximately equal contrast. Image intensities are locally adjusted by estimating gamma correction and contrast adjustment parameters in a number of image regions. Interpolating these parameters for every image pixel yields smooth parameter surfaces that are capable to correct illumination and light scattering variations of regional scales.

### 3.3.3 Snow Detection

Snow detection was found to be a crucial part, when monitoring permafrost surface creep over a long period of time, i.e., over several months at its minimum. The snow-free period typically is relatively short as these objects are located at high altitudes and usually face the northern side of the mountain slopes. Specially during the snow melting period, there are large areas free of snow next to snow filled depression regions that take a long time to become free of snow. Fig 3.6 shows the effect of a melting snow patch over the course of four consecutive days. As described before (Section 3.3.1), image features for matching are extracted automatically based on appropriate image contrast. The transition area between image pixels representing snow and those showing rocks has an exceptionally high contrast such that it is very likely that good feature points are defined along the snow-rock boarder. As shown in Section 3.2.1, the driving force for the non-linear adjustment of feature matching between images is the image gradient. Thus for these features the transformation being estimated will not show the motion of rocks but the retreat of snow as it melts.

Areas covered by snow have to be detected prior to the feature selection and feature matching process. The accuracy of this detection, however, does not need to be extremely accurate, as long as prominent snow patches are well discovered. Smaller patches are generally detected as gross errors during the matching process, as long as their relative influence remains below the detection threshold (typically not more than 40% - very robust implementations are near the 50% border (*Fellbaum*, 1994)).

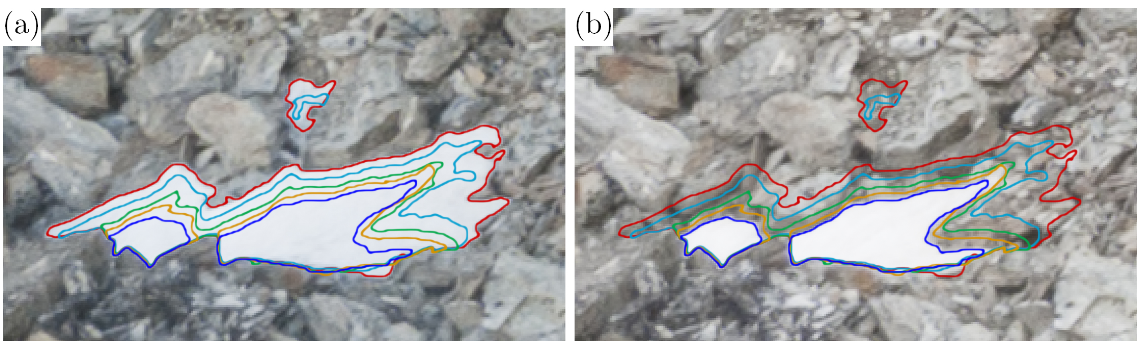


**Figure 3.6:** Melting patch of snow monitored over five consecutive days in June 2014. The background image in (a) is for day 1 and in (b) for day 5. Contours show the changing snow patch extension.

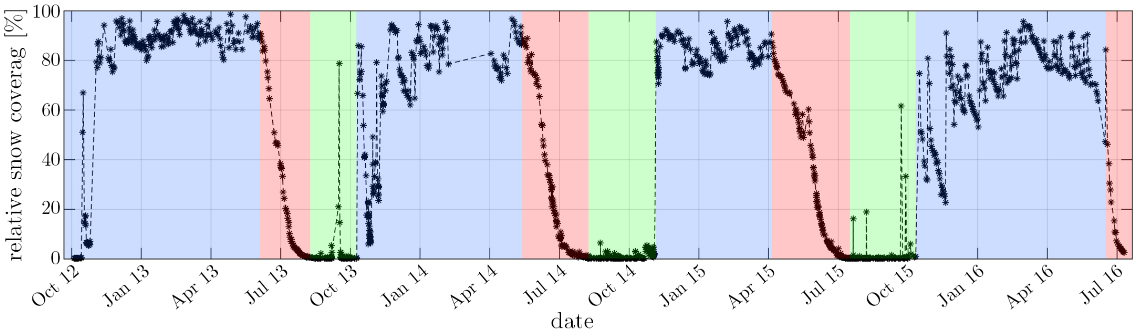


**Figure 3.7:** Relative snow coverage seen from camera position 2 over the image acquisition time span. Red areas indicate the snow melt period, defined by the continuous retreat of snow, Areas in green and blue show the periods without snow and full approximately full snow coverage, respectively.

Snow detection in digital images was and still is an active area of research (e.g. *MacQueen*, 1967; *Rüfenacht et al.*, 2014). As imaging conditions are prone to strong variations (changing illumination, shadow casting, exposure mismatch, white balance issues, etc.), snow detection by simple intensity clipping is not a feasible method and snow detection is not as easy as it may seem. As the accuracy does not need to be very high, the approach used in this work is based on analyzing different image segments that are obtained using a k-means clustering procedure. Generally, snow patches have three characteristics in common: high luminosity, low image gradient within the snow areas, and a strong image gradient along the snow-rock boundary. These three quantities are defined for each cluster type and used to identify clusters representing snow. Three parameters (one for each criterion) were empirically determined using a couple of images taken under different conditions. Pixels detected as snow are not considered for matching or any other analysis during the image processing. Regarding the matching process, snow pixels within the matching window are treated as outliers a priori by setting the corresponding entries in the weight matrix $\mathbf{Q}_{hh}^{-1}$ (Eqn. (3.9)) to zero.

Fig 3.7 shows the number of detected snow pixels with respect to the total number of pixels defined for a given reference area [1] over the course of nearly four years. It is clear that snow detection not only has to detect snow but also to declare an image to be free of snow if there is non. A few examples of detected snow areas are given in Appendix A.

[1] counted using a mask defining the extent of the reference area in Fig. A.1, Appendix A

# 4 Object Point Reconstruction

In geodesy, remote sensing, and a variety of other applications, one of the key tasks is to estimate 3-dimensional coordinates from points in space (e.g. *Mayer*, 1999). Given the problem statement, different methods exist - GPS being the most well-known of them. As stated in Chapter 1, the project aims at low-cost monitoring using a combination of low-cost GNSS and optical camera devices. This chapter will give the theoretical background of 3-dimensional reconstruction using optical images as observations of a scene. Each photograph shows the 3D scene projected into the 2D image. By combining a minimum of two photographs taken from the same scene but with a different viewing angle, the principals of scene reconstruction can be applied. Sometimes in geomonitoring, only one camera is used to estimate 3D coordinates (e.g. *Travelletti et al.*, 2012). For such scenarios, a complementary dataset, typically a Digital Elevation Model (DEM), is needed.

The first section of this chapter introduces the coordinate system used in this study. Section 4.2 then gives an overview about the principles and problems of object point reconstruction using a single camera and a DEM. In Section 4.3, the mathematical principles of 3-dimensional coordinate estimation based on at least two views is explained. A more detailed theory on the estimation procedure using the principle of bundle adjustment is explained in Section 4.4. The problem of initial parameter estimates for bundle adjustment is explained in Section 4.5 and finally, in Section 4.6, the combination principle of the photogrammetric reconstruction process with GNSS position estimates is explained.

## 4.1 Coordinate System and Conventions

Before the fundamental equations can be introduced, the coordinate system, wherein the computations take place, has to be defined. There are different definitions found in literature: *Hartley and Zisserman* (2003) define a right-handed coordinate system for world and camera coordinates, where the camera looks into the positive z-direction. *Kraus* (2007) also uses a right-handed frame, but the camera looks into the negative z-direction. This work uses the latter convention.

Typically, the origin of the image or frame coordinate system is located at the image center. The 2-dimensional coordinate system, defined by image axes $x'$ and $y'$ is extended by an axis $z'$ orthogonal to the image plane. The origin of this 3D coordinate system is located at the perspective center $O$. An image is produced by projecting rays through the perspective center, building a negative image on the sensor. For mathematical reasons it is more convenient to work with the positive image (see Fig. 4.1), thus defining the principle distance $c$ in the negative z-direction (*Luhmann et al.*, 2006).

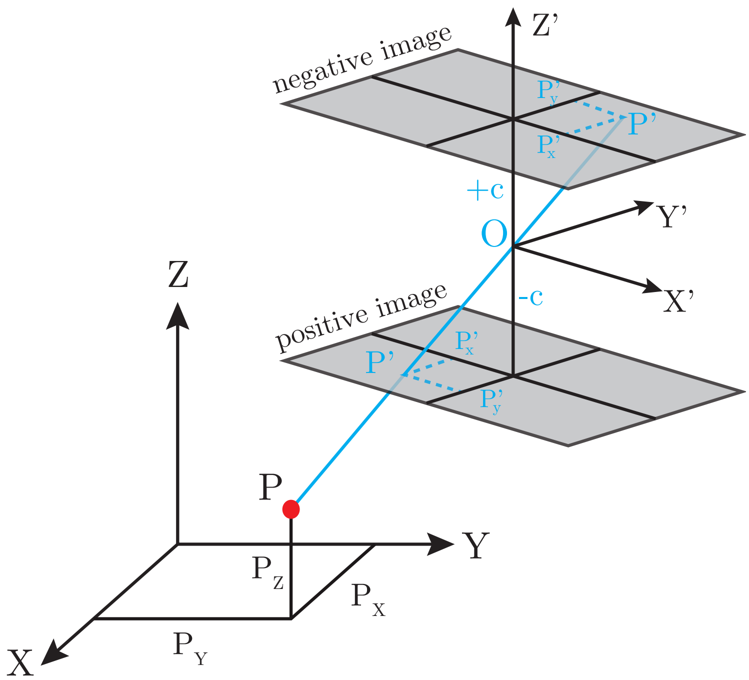


**Figure 4.1:** The world (X, Y, Z) and camera (X', Y', Z') coordinate system used in this work. The camera looks into the negative z-direction. As the positive image is used for mathematical derivations, the principle distance is $-c$.

The world (or object) coordinate system, also a right-handed system, defines $X$ = easting, $Y$ = northing, and $Z$ = altitude.

### 4.1.1 Topocentric Cartesian System

Usually geodetic coordinates are given in a national coordinate system. For example digital elevation models for Switzerland are computed in the reference frame LV03 or LV95. These systems use ellipsoidal heights that do not correspond to a Cartesian coordinate frame used and assumed to hold for the reconstruction process. In case of small spatial regions of interest, the difference between the non-Cartesian and Cartesian frame is negligible. When measuring object coordinates over 100 meter and more, the difference between the two reference frames is notable. I.e., for 100 m the coordinate difference can be as high as 47 cm for the $X-$ or $Y-$ and around 8 mm for the height component. It is thus important to work in a fully Cartesian coordinate system. If the original non-Cartesian frame is used, the systematic model error needs to be corrected by appropriate correction terms (e.g. *Gruen*, 1986).

In this work, the DEM coordinates, the ground control points, as well as the GNSS coordinates for the stations in the FOV are converted into a geocentric coordinate system, where the initial position of camera number 1 (see Chapter 6) is used as the topocentric origin.

### 4.1.2 Rotation in Space

Rotation in space is treated as a combination of three rotations around the individual Cartesian axes, with Eulerian angles $\omega$, $\phi$, and $\kappa$:

$$\mathbf{R} = \mathbf{R}_x(\omega) \cdot \mathbf{R}_y(\phi) \cdot \mathbf{R}_z(\kappa) \tag{4.1}$$

$$= \begin{pmatrix} 1 & 0 & 0 \\ 0 & c_\omega & -s_\omega \\ 0 & s_\omega & c_\omega \end{pmatrix} \cdot \begin{pmatrix} c_\phi & 0 & s_\phi \\ 0 & 1 & 0 \\ -s_\phi & 0 & c_\phi \end{pmatrix} \cdot \begin{pmatrix} c_\kappa & -s_\kappa & 0 \\ s_\kappa & c_\kappa & 0 \\ 0 & 0 & 1 \end{pmatrix} \tag{4.2}$$

$$= \begin{pmatrix} c_\phi s_\kappa & -c_\phi s_\kappa & s_\phi \\ c_\omega s_\kappa + s_\omega s_\phi c_\kappa & c_\omega c_\kappa - s_\omega s_\phi s_\kappa & -s_\omega c_\phi \\ s_\omega s_\kappa - c_\omega s_\phi c_\kappa & s_\omega c_\kappa + c_\omega s_\phi s_\kappa & c_\omega c_\phi \end{pmatrix} \tag{4.3}$$

with $c_\phi = \cos(\phi)$, $s_\omega = \sin(\omega)$, etc. The primary rotation is defined around the x-axis, the secondary around the y-axis and the tertiary around the z-axis.

## 4.2 Single-View

Points in space can be reconstructed using a single camera, in the following called single-view (vs. multi-view for several cameras), combined with a DEM that is used as a projection surface (Fig. 4.2). In literature, this technique is known as monoplotting. Decades ago, *Gruen and Sauermann* (1977) proposed this method with amateur cameras.

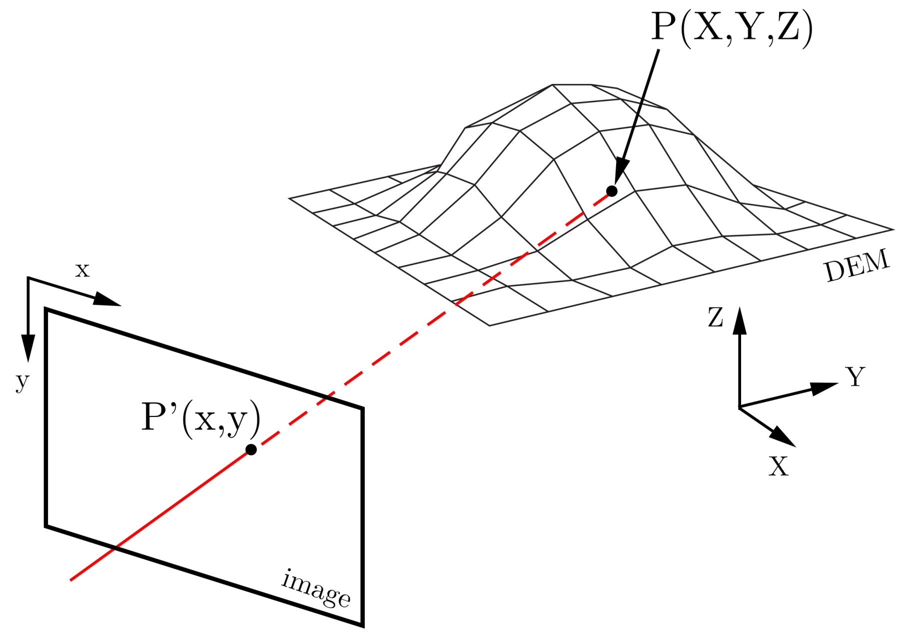


**Figure 4.2:** The principle of monoplotting. A single camera, oriented in space, combined with a DEM is used to estimate a 3D position of a measured point in the 2-dimensional image space.

When using a DEM as the projection surface, errors introduced by the elevation model are usually large compared to the required accuracy of the object point estimation (*Fisher and Tate*, 2006), specially if the model is not up-to-date and the surface is expected to change in time. This type of data merging is often used in satellite remote sensing, where the estimated surface displacements are projection onto a DEM to obtain the ground geometry and to scale the estimated displacements from relative to absolute units (e.g. *Berthier et al.*, 2005).

### 4.2.1 DEM Intersection

If the camera parameters are known (Section 4.5.3), image coordinates can be used for raytracing to find the intersecting points on a DEM. To do so, *Möller and Trumbore* (1997) proposed an algorithm to efficiently intersect rays with planes that are defined by three points in space: $A$, $B$, and $C$. The intersection point $P$ can then be expressed as:

$$\mathbf{p} = \mathbf{a} + u \cdot (\mathbf{b} - \mathbf{a}) + v \cdot (\mathbf{c} - \mathbf{a}) \tag{4.4}$$

The same point $\mathbf{p}$ can also be defined using the parametrized ray equation with origin $\mathbf{o}$ and direction $\mathbf{d}$:

$$\mathbf{p} = \mathbf{o} + t \cdot \mathbf{d} \tag{4.5}$$

An exact solution of point $P$ can then be found by solving for $t$, $u$, and $v$:

$$\begin{pmatrix} t \\ u \\ v \end{pmatrix} = \begin{pmatrix} -\mathbf{d} & (\mathbf{b} - \mathbf{a}) & (\mathbf{c} - \mathbf{a}) \end{pmatrix}^{-1} \cdot \begin{pmatrix} \mathbf{o} - \mathbf{a} \end{pmatrix} \tag{4.6}$$

When computing $\mathbf{p}$ by Eqn. 4.5 or 4.4, the cofactor matrix of the intersection point is found by error propagation:

$$\mathbf{K}_{PP} = \mathbf{F}^T \cdot \mathbf{K}_{ll} \cdot \mathbf{F} \tag{4.7}$$

$$\text{with} \quad \mathbf{F}^T = \begin{pmatrix} \frac{\partial p_x}{\partial l_1} & \frac{\partial p_x}{\partial l_2} & \cdots & \frac{\partial p_x}{\partial l_{27}} \\ \frac{\partial p_y}{\partial l_1} & & \ddots & \vdots \\ \frac{\partial p_z}{\partial l_1} & & \cdots & \frac{\partial p_z}{\partial l_{27}} \end{pmatrix} \tag{4.8}$$

Here 27 parameters are involved: 16 camera parameters (Section 4.5.3), 2 image coordinates, and 9 components for the three object points $\mathbf{a}$, $\mathbf{b}$, and $\mathbf{c}$. Typically, the elevation data in a DEM is given on a regularly spaced grid of East and North components, thus the horizontal position errors can be neglected, whereas the height component errors heavily depend on the resolution, the slope steepness, and the method the DEM was generated with (*Fisher and Tate*, 2006).

A simple example of error propagation for a ray-plane intersection at different incident angles is shwon in Fig. 4.3. Table 4.1 lists the parameters used in the example. More details about the meaning of the individual (image) parameters is given in Section 4.3.2.

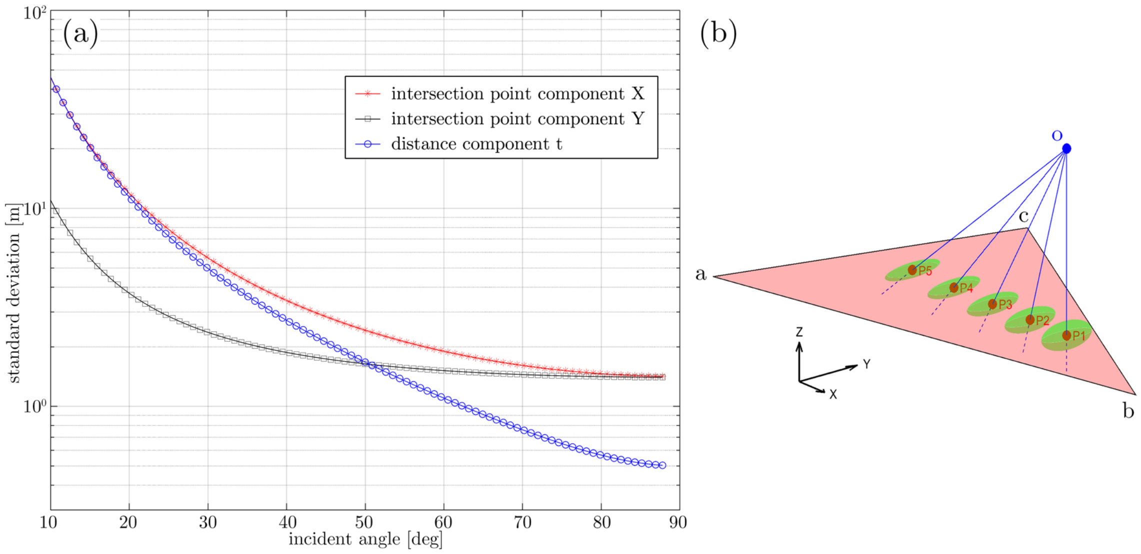


**Figure 4.3:** (a) Effect of ray-plane intersection angle and the estimated error of the intersection point $P$ (components X and Y). Also given are error estimates of the distance between the camera center and point $\mathbf{p}$. The parameters used are given in Table 4.1. $\sigma_{\mathbf{p}_Z} = 0$, as points **a**, **b**, and **c** defining the plane are assumed to be free of error. (b) Test geometry of rays with origin **o**, intersecting with the plane defined by the points **a**, **b**, and **c**. Shaded in green are the covariance ellipsoids.

As the points **a**, **b**, and **c** are given with zero error, the $Z$-component of the intersection point has no error either ($\sigma_{p_Z} = 0$). Thus, only the errors of the $X$, and $Y$ component as well as the error of the distance measure, variable $t$, are shown. For intersection angles

**Table 4.1:** Parameters used to compute the ray-plane intersection error. Covariance components were all zero. The camera orientation was set to $\omega = 12^\circ$, $\kappa = 0^\circ$, and $\phi$ was iteratively changed from 10 to 90°, all with $\sigma_{\omega,\kappa,\phi} = 0.57^\circ$.

| $l_{\text{image}}$ | value $[mm]$ | $\sigma$ $[mm]$ | $l_{\text{objects}}$ | value $[m]$ | $\sigma$ $[m]$ |
|---|---|---|---|---|---|
| $x_i$ | 0.5 | 0.1 | $O_X$ | 12.0 | 0.2 |
| $y_i$ | −2.0 | 0.1 | $O_Y$ | 0.5 | 0.2 |
| $x_p$ | 0.1 | 0.01 | $O_Z$ | 109.0 | 0.5 |
| $y_p$ | 0.2 | 0.01 | $A_X$ | −1.0 | 0.0 |
| $c$ | 10.0 | 0.005 | $A_Y$ | −2.0 | 0.0 |
| $k_1$ | $1.0e^{-4}$ | $1.0e^{-6}$ | $A_Z$ | 9.0 | 0.0 |
| $k_2$ | 0.0 | $1.0e^{-7}$ | $B_X$ | 20.0 | 0.0 |
| $k_3$ | 0.0 | $1.0e^{-9}$ | $B_Y$ | 0.5 | 0.0 |
| $p_1$ | 0.0 | $1.0e^{-10}$ | $B_Z$ | 9.0 | 0.0 |
| $p_2$ | 0.0 | $1.0e^{-10}$ | $C_X$ | 0.0 | 0.0 |
| $sc$ | 0.0 | 0.0 | $C_Y$ | 6.0 | 0.0 |
| $sh$ | 0.0 | 0.0 | $C_Z$ | 9.0 | 0.0 |

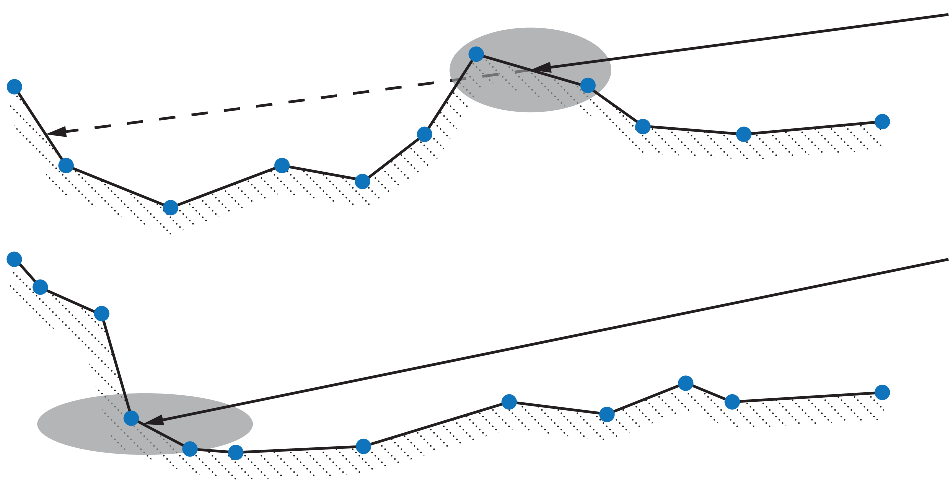

**Figure 4.4:** Two examples of ray surface intersections shown in cross section. Blue dots show the DEM heights and the grey ellipses represent the formal error of the intersection point. The top illustration shows a critical scenario of error underestimation, whereas the bottom scenario yields an overestimation of the error.

close to 90 degrees, the error in $t$ is dominated by the camera position and for the $X$ and $Y$ components by the intrinsic camera parameters (left part in Table 4.1). With decreasing incident angle, the error of all components increases, while the error in $t$ increases fastest and asymptotically approaches the error of the $X$-component. The latter behavior is observed, as the incident angle is modified by a rotation around the $y$-coordinate axis ($\phi$).

### 4.2.2 Monte Carlo Simulation for Error Estimation

Error propagation as shown in the previous Section (4.2.1) is valid only, if the plane described by the three points **a**, **b**, and **c** is infinite. When rays are intersected with a DEM, however, the valid extent of the intersection plane is defined by the three points of the triangle. As neighboring planes have different orientations, other errors result, if a ray misses the expected intersection plane. Two examples showing scenarios where the correctness of the formal error is under- and over-estimated, are shown in Fig. 4.4. The problem cannot be solved analytically because the distribution of the intersection parameters **p** is not Gaussian anymore.

To overcome the problem of formal error estimation, an empirical error analysis using the Monte Carlo simulation method is applied. The principle here is to create realizations of rays (Eqn. (4.5)) following the relative frequencies of probabilities given by the covariance matrix of the parameters. Doing so, the correlation of the random variables is considered such that each realization follows a multivariate normal distribution (more details on the realization of this procedure is given in Section 5.2).

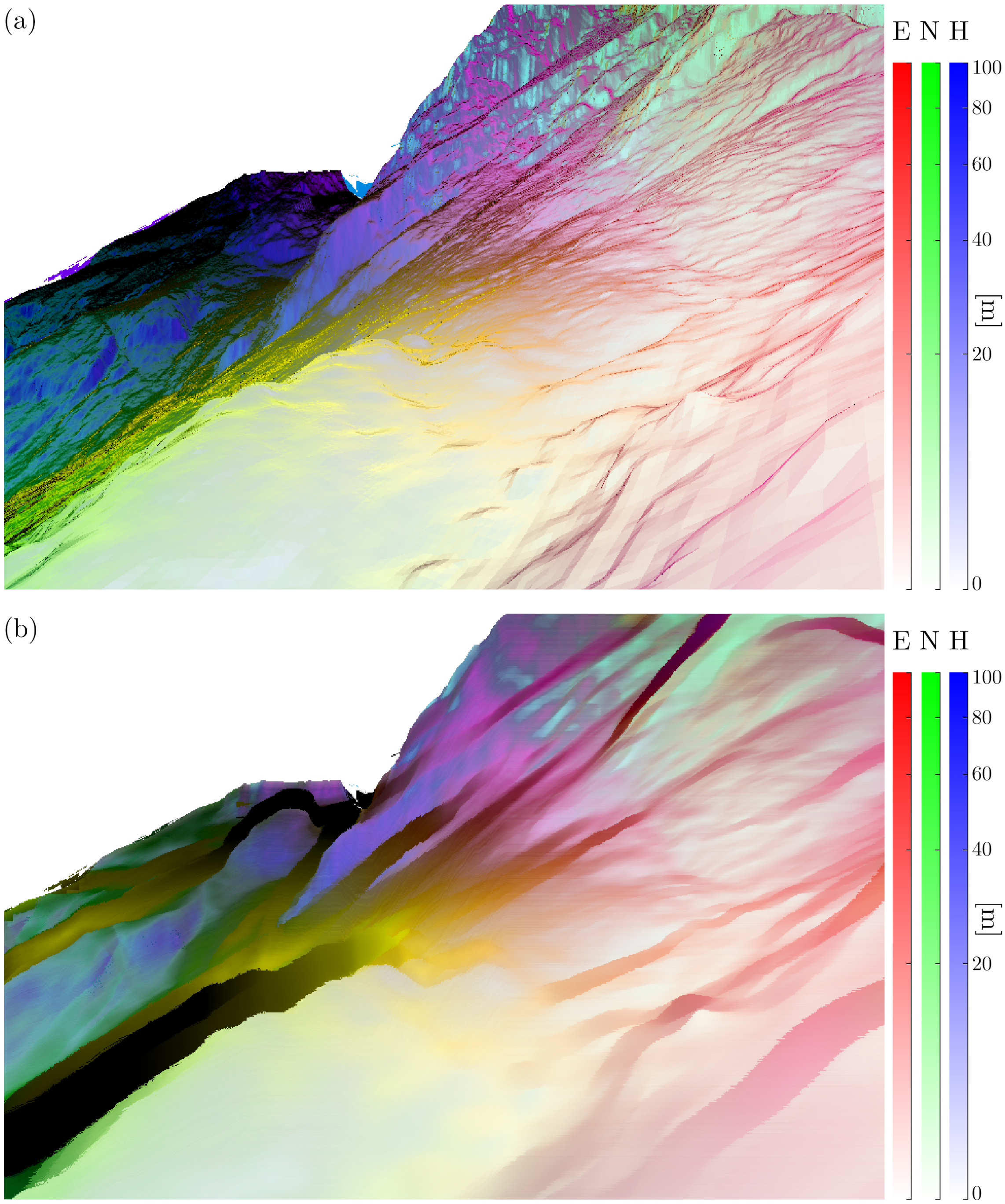


**Figure 4.5:** (a) Difference of 95% and 5% quantiles obtained by formal error propagation for ray tracing using a $1\,m$ resolution DEM (provided by the Federal Office of Environment). (b) 95% and 5% quantile difference of the Monte Carlo simulation result. Colors define the respective components (East, North, and Height) for the intersection points. Both figures have the same logarithmic scale.

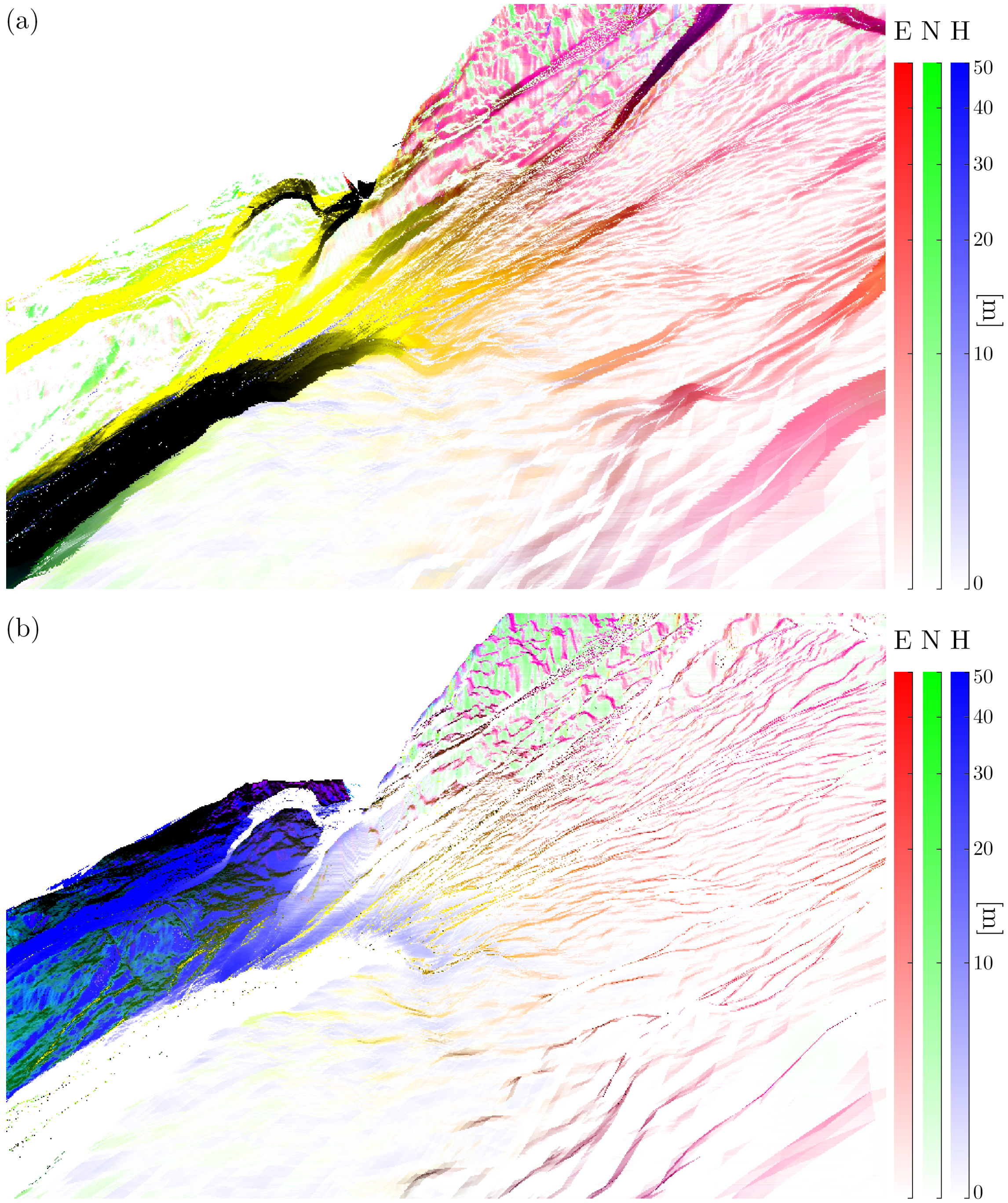


**Figure 4.6:** Differences between the formal error propagation and Monte Carlo simulation: (a) areas, where the formal error is too low, and (b) where the formal error is too high. Both figures have the same logarithmic scale as also in Fig. 4.5, however, the color scale spans only half the range. Dark areas are beyond the color scale (too large).

For each intersection bundle, a different distribution is obtained, so that the standard deviation is not an adequate quantity to be compared with the formal error estimates. As an alternative, the difference between the 95% and 5% quantiles are used for comparison. Fig. 4.5 shows the estimated quantile differences for the formal error as well as for the Monte Carlo simulation. The camera azimuth orientation is close to 45 degrees, leading to near-field errors that are dominated by the North component in the left and by the East component in the right image area. Steep rock walls (e.g., near the middle of the frame) are dominated by errors in the vertical (Height) component and thus appear in blue. For zones above the $100\,m$ quantile difference, luminosity decreases continuously.
Comparing the two results (Fig. 4.5) clearly shows that there are areas with heavy error underestimation and areas, where the formal error is too high, more clearly shown in Fig 4.6. For distant regions, error estimation using the Monte Carlo simulation yields smooth transitions between the different zones. Because small differences in simulated ray directions yield a larger dispersion of planes being intersected, a smoothing effect results for the estimated errors between neighboring ray bundles.

Both, the Monte Carlo and the formal error estimates were computed with zero error in the DEM components, thus shows an idealized scenario. Given the different error sources of a DEM along with possible temporal inconsistencies, reconstruction of objects in space using this sort of data bears the potential of erroneous error estimation.

### 4.2.3 Image Recification

Image rectification defines a transformation of image coordinates into a specific image plane (i.e., a reference coordinate system defined w.r.t. a specific image). There are different possibilities to estimate the parameters for this transformation. In the following, a short overview of three possibilities to conduct this transformation are presented. Note that some definitions used here are described in more detail in the following sections of this chapter.

#### Epipolar Geometry

Generally, the problem can be solved by the principles of Epipolar Geometry (e.g. *Hartley and Zisserman*, 2003). In a general case for cameras with unknown intrinsic parameters (see Section 4.5.1), the Fundamental matrix $\mathbf{F}$ relates object points in space seen from two views by nine parameters given in the $3 \times 3$ matrix $\mathbf{F}$. In case of image coordinates $\mathbf{x}_{1,2}$ of camera 1 and 2, respectively, the essential matrix $\mathbf{E}$ can be estimated instead. Mathematically, the following relationship is given (with $\mathbf{x}_1$ and $\mathbf{x}_2$ in homogeneous coordinates):

$$\mathbf{x}_1^T \cdot \mathbf{F} \cdot \mathbf{x}_2 = 0 \tag{4.9}$$
$$\mathbf{x}_1^T \cdot \mathbf{E} \cdot \mathbf{x}_2 = 0 \tag{4.10}$$
$$\mathbf{E} = \mathbf{K}_1^T \cdot \mathbf{F} \cdot \mathbf{K}_2 \tag{4.11}$$

with $\mathbf{K}_{1,2}$ represent the camera calibration matrices, see also Eqn. (4.52).

### 2D Transformation

A two dimensional transformation as presented in Section 3.2 can be estimated for the whole image to correct for example for a simple translation, a similarity or an affine transformation. The model can be extended, e.g., to include additional correction parameters for lens distortion (Eqn. (4.15)). This type of transformation can, e.g., be applied when a distant scene is observed from the same position in space. In case the latter condition is not fulfilled, the mathematical model is not valid (as the projection follows the collinearity principles, Eqn. (4.14)).

### Known Camera Parameters

The third method presented here shows the principle of image rectification in case the intrinsic and extrinsic camera parameters are known (i.e., a set of 16 parameters for both cameras), see Section 4.3.2. For the purpose of ray tracing, the reference image coordinates are chosen to be the corrected (i.e., calibrated) image coordinates ($x_{\text{corr}}$,$y_{\text{corr}}$), see Eqn. (4.15). Given a reference image, the transformation into the corrected coordinates of the reference image can be obtained in a three step procedure:

- Correction of image coordinates using the camera parameters of the image to be transformed (Eqn. (4.15)).
- Projection of image coordinates into the 3D space using the ray Eqn. (4.5).
- Back-projection of the space coordinates using the collinearity equations with parameters corresponding to the reference image (Eqn. (4.14)).

Mathematically this rectification process can be written as:

$$x_{\text{ref}} = -c_{\text{ref}} \frac{\mathbf{r}_{1r} \cdot \Delta\mathbf{m}}{\mathbf{r}_{3r} \cdot \Delta\mathbf{m}} \tag{4.12}$$

$$y_{\text{ref}} = -c_{\text{ref}} \frac{\mathbf{r}_{2r} \cdot \Delta\mathbf{m}}{\mathbf{r}_{3r} \cdot \Delta\mathbf{m}} \tag{4.13}$$

$$\text{whereas} \quad \Delta\mathbf{m} = \begin{pmatrix} O_X - O_{X_{\text{ref}}} + \mathbf{r}_1 \cdot \mathbf{d} \\ O_Y - O_{Y_{\text{ref}}} + \mathbf{r}_2 \cdot \mathbf{d} \\ O_Z - O_{Z_{\text{ref}}} + \mathbf{r}_3 \cdot \mathbf{d} \end{pmatrix} = \begin{pmatrix} \Delta O_X + \mathbf{r}_1 \cdot \mathbf{d} \\ \Delta O_Y + \mathbf{r}_2 \cdot \mathbf{d} \\ \Delta O_Z + \mathbf{r}_3 \cdot \mathbf{d} \end{pmatrix}$$

$$\text{and} \quad \mathbf{R}_{\text{ref}} = \begin{pmatrix} \mathbf{r}_{1r}^T & \mathbf{r}_{2r}^T & \mathbf{r}_{3r}^T \end{pmatrix}$$

$$\mathbf{R} = \begin{pmatrix} \mathbf{r}_1 \\ \mathbf{r}_2 \\ \mathbf{r}_3 \end{pmatrix}$$

$$\mathbf{d} = \begin{pmatrix} x_{\text{corr}} & y_{\text{corr}} & -c \end{pmatrix}^T$$

with $(..)_{\text{ref}}$ indicating the parameters associated with the reference epoch, others are for the image to be transformed.

## 4.3 Stereo-/Multi-View

### 4.3.1 Reconstruction Principle

In this section, the theory of three-dimensional reconstruction using two or more cameras is given. This type of object reconstruction is widely used (e.g. *Remondino and El-Hakim*, 2006) and well studied as no additional information other than camera-related observations are used for the reconstruction process. To define the coordinate system, i.e., the geodetic datum, also (ground) control point coordinates are needed, see Fig. 4.7. In applications like the monitoring of permafrost creep or similar phenomena, where only a limited number of cameras can be used to observe large spatial areas, spatial intersection (discussed in Section 4.4) may yield good results only for a small patch in the overlapping images.

The principle of image-based reconstruction using two or more images relies on the recognition of individual objects between the views. Once a number of such relations has been found, the camera positions, orientations, as well as other parameters needed to reconstruct the coordinates of the objects of interest, can be estimated. The aim of the following section is to give an overview of the mathematical background used for this purpose.

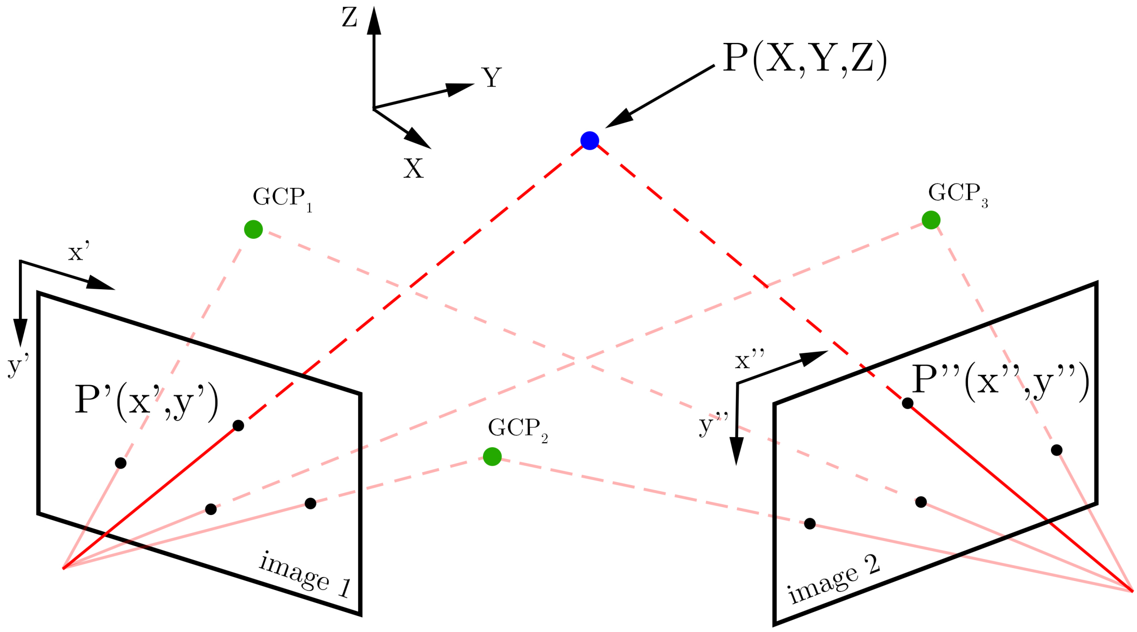


**Figure 4.7:** The principle of stereo reconstruction. At least two cameras and a set of ground control points (GCP, in green) are used to find all parameters included in the estimation of a 3D position of a point **P** seen in both views.

### 4.3.2 Mathematical Model

When dealing with photogrammetric scene reconstruction, it is important to use an adequate mathematical model to describe the effects involved, when a scene point is projected onto the image plane, yielding an image coordinate as primary measurement information. The principal model of scene projection can be described using geometric relations in the pinhole camera model (e.g. *Luhmann et al.*, 2006; *Kraus*, 2007):

$$\begin{aligned} x &= x_p - c \cdot \frac{r_{11}(X-X_0)+r_{21}(Y-Y_0)+r_{31}(Z-Z_0)}{r_{13}(X-X_0)+r_{23}(Y-Y_0)+r_{33}(Z-Z_0)} = x_p - c \cdot \frac{M_1}{M_3} \\ y &= y_p - c \cdot \frac{r_{21}(X-X_0)+r_{22}(Y-Y_0)+r_{32}(Z-Z_0)}{r_{13}(X-X_0)+r_{23}(Y-Y_0)+r_{33}(Z-Z_0)} = y_p - c \cdot \frac{M_2}{M_3} \end{aligned} \tag{4.14}$$

where the symbols are:

$x, y$ … metric image x,y-coordinates [mm]
$x_p, y_p$ … image coordinates of the principal point [mm]
$c$ … principal distance [mm]
$r_{ij}$ … elements of the rotation matrix $\mathbf{R}$ (Eqn. 4.3)
$X, Y, Z$ … cartesian object coordinates in space [m]
$X_0, Y_0, Z_0$ … cartesian coordinates of the camera center in space [m]

The derivation of these collinearity equations can be found in many textbooks (e.g. *Kraus*, 2007) and they are the very principle of projective geometry. As any camera differs from such an ideal pinhole camera, the projective rays are affected by refraction and imperfections in the manufactured lenses in use (e.g. *Clarke and Fryer*, 1998). *Brown* (1971) has shown that most camera lens distortion effects can well be modeled by five additional parameters (AP) $k_1$, $k_2$, $k_3$, $p_1$, and $p_2$. In rare cases of different scaling in $x$ and $y$ or a skew effect between these coordinate axes two additional parameters ($s_c$ and $s_h$) can account for it. The equations for the corrected image coordinates ($x_{corr}$, $y_{corr}$) are then given by:

$$\begin{aligned} x_{corr} &= x_c \cdot (1+dr) + p_1 \cdot (r^2 + 2 \cdot x_c^2) + 2p_2 \cdot x_c y_c - s_c \cdot x_0 + s_h \cdot y_0 \\ y_{corr} &= y_c \cdot (1+dr) + 2p_1 \cdot x_c y_c + p_2 \cdot (r^2 + 2 \cdot y_c^2) + s_h \cdot x_0 \end{aligned} \tag{4.15}$$

with

$$\begin{aligned} x_c &= x - x_p \\ y_c &= y - y_p \\ r^2 &= x_c^2 + y_c^2 \\ dr &= k_1 \cdot r^2 + k_2 \cdot r^4 + k_3 \cdot r^6 \end{aligned} \tag{4.16}$$

Considering the light travelling path of objects projecting in image areas with the same radial distance, $r$ (Eqn. (4.16)), *Fraser and Shortis* (1992) have shown that distortion varies according to the object distance and increases with the image scale (decreasing object distance). *Dold* (1997) proposed a distance-dependent distortion correction term that requires, however, a very strong network in order to yield reliable parameters. As the image scale in this study does not vary much and as there are only two cameras located at permanent positions (Chapter 2), the network is relatively weak and thus the additional distance-dependent distortion parameters cannot be reliably determined. Additional notes on lens distortion can be found in Section 4.4.6.

The collinearity equations (4.14) describe the projection of object points into a perfect, i.e., undistorted image with a principle point offset $x_p$, $y_p$, or mathematically:

$$\begin{aligned} x_{corr} &= -c \cdot \frac{M_1}{M_3} \\ y_{corr} &= -c \cdot \frac{M_2}{M_3} \end{aligned} \tag{4.17}$$

The combination of the collinearity equations with the correction terms, Eqn. 4.15, then results in:

$$\begin{aligned} 0 &= c \cdot \frac{M_1}{M_3} + x_c \cdot (1 + dr) + p_1 \cdot (r^2 + 2 \cdot x_c^2) + 2p_2 \cdot x_c y_c - s_c \cdot x_0 + s_h \cdot y_0 \\ 0 &= c \cdot \frac{M_2}{M_3} + y_c \cdot (1 + dr) + 2p_1 \cdot x_c y_c + p_2 \cdot (r^2 + 2 \cdot y_c^2) + s_h \cdot x_0 \end{aligned} \tag{4.18}$$

Often the scale and skew coefficients $s_c$ and $s_h$ do not differ significantly from zero (see Appendix B). For projects, where image acquisition takes place in a relatively stable environment (over the measurement period), distortion parameters are estimated in a calibration procedure taking place just before or after the photographs have been taken (see Section 4.5.3). The advantage hereby is, that 3D object points can be reconstructed in a relatively simple way, as these coordinates (along with the positions of the camera centers and the orientation angles) are the only unknowns to be solved for. In contrast to such a procedure, there are situations, where camera parameters need to be estimated during the optimization of the model (e.g. *Pedersini et al.*, 1998). These problems are solved using the well-known bundle adjustment method[1], where the residuals of the non-linear Eqn. (4.18) are minimized, solving for all significant parameters. Although the problem of bundle adjustment is well-known since many decades (*Brown*, 1976), it is still not an easy task to make it efficient and robust for all possible scenarios (*Triggs et al.*, 2000). One critical issue, when estimating 3D object coordinates using bundle adjustment, is that the highly non-linear system (4.18) needs to be initialized with approximate values (see Section 4.5.1).

There are different strategies to estimate 3D object coordinates, when there are also camera parameters that need optimization. Doing a full bundle adjustment is the proper way, specially regarding statistical aspects (i.e., correlation between the various parameters). The main problem here is, that there are $3 \times N + 16 \times M$ parameters to be estimated ($N$ is the number of object points, $M$ the number of cameras involved), thus the matrix to be inverted easily gets very large. Indeed, there are many software tools available (a few of them mentioned in *Triggs et al.* (2000)) that are specifically optimized for such large problems. The following Section (4.4) is dedicated to the bundle adjustment problem, including its statistical evaluation.

[1]Bundle adjustment has its name from the fact that bundle of rays (projected through the image centers) are used for estimation.

## 4.4 Bundle Adjustment

### 4.4.1 Principles

The principle of bundle adjustment found in the literature is formulated as the following Gauss-Markov system (e.g., *Brown*, 1976; *Gruen*, 1986; *Triggs et al.*, 2000; *Hartley and Zisserman*, 2003):

$$\begin{pmatrix} \mathbf{A}_n & \mathbf{A}_g & \mathbf{A}_e & \mathbf{A}_a \\ & \mathbf{I} & & \\ & & & \mathbf{I} \end{pmatrix} \cdot \begin{pmatrix} \mathbf{p}_n \\ \mathbf{p}_g \\ \mathbf{p}_e \\ \mathbf{p}_a \end{pmatrix} + \begin{pmatrix} \mathbf{v}_b \\ \mathbf{v}_g \\ \mathbf{v}_a \end{pmatrix} = \begin{pmatrix} \mathbf{l}_b \\ \mathbf{l}_g \\ \mathbf{l}_a \end{pmatrix} \tag{4.19}$$

with subscripts $n$ ... new object point coordinates
$g$ ... ground control point coordinates
$e$ ... exterior camera parameters (for each view)
$a$ ... additional camera parameters (for each view)
$\mathbf{A}_n \dots \mathbf{A}_a$ ... the corresponding coefficient matrices
$\mathbf{p}_n \dots \mathbf{p}_a$ ... the corresponding parameter vectors

As the primary observation equations (4.19) are defined by Eqn. 4.18, the coefficient matrices $\mathbf{A}_n \dots \mathbf{A}_a$ contain the respective partial derivatives for each parameter involved[2].

The observations are given by the image point coordinates $\mathbf{l}_b$ (with corresponding residuals $\mathbf{v}_b$), the pseudo-observations of the ground control points $\mathbf{l}_g$ (with corresponding residuals $\mathbf{v}_g$), and the pseudo-observations of the additional parameters $\mathbf{l}_a$ (with corresponding residuals $\mathbf{v}_a$). For the stochastic model, the following relation holds:

$$\mathbf{v}_b \sim \mathcal{N}(0; \mathbf{K}_{bb}) \tag{4.20a}$$
$$\mathbf{v}_g \sim \mathcal{N}(0; \mathbf{K}_{gg}) \tag{4.20b}$$
$$\mathbf{v}_a \sim \mathcal{N}(0; \mathbf{K}_{aa}) \tag{4.20c}$$

The matrices $\mathbf{K}_{bb}$, $\mathbf{K}_{gg}$, and $\mathbf{K}_{aa}$, being the respective variance-covariance matrices of the observations. The additional pseudo-observations for ground control points define the geodetic datum. Also the additional parameters (AP) are traditionally added as pseudo observations (e.g. *Gruen*, 1986), as these are helpful for:

- Stabilizing the inversion of the normal matrix, also allowing the parameters not to take an arbitrary value but to move slightly from their original guesses, $l_{a,i}$.
- A flexible integration of real observations of the additional parameters (e.g., from external calibration)

[2] Note that in a very strict formulation, the system has to be written in the form $\mathbf{f}(\check{l}, \check{p}) = 0$ rather than $\mathbf{f}(\check{p}) = \check{l}$, as shown in Eqn. C.4 (Appendix C)). Neglecting the additional partial derivatives with respect to $\check{l}$, however, is feasible as they have a vanishing effect on the weight matrix used to build the parameter covariance matrix. As this is an iterative approach, no difference will be observed in the final parameter estimates and its corresponding cofactor matrix.

- Constraining a non-significant parameter $l_{a,i}$ (found during the adjustment process - Section 4.4.2 - or from a priori information) by setting $l_{a,i} = 0$ and $\mathbf{K}_{a_i a_i} \rightarrow 0$.
- Building a very flexible system, where parameters can also be set completely free by $\mathbf{K}_{a_i a_i} \rightarrow \inf$.

The strength of the formulation 4.19 is that it allows to quickly adjust to a given geometry and a priori information, which is a very valuable characteristic. The same implementation can, e.g., also be used to make a pure estimation of camera parameters based on fixed object coordinates for a single camera. On the other hand, pure spatial intersection can be performed, when constraining the camera parameters with pseudo-observations using: $\mathbf{K}_{ll,e,1} = \mathbf{K}_{ll,e,2} = \ldots \mathbf{K}_{ll,e,v} \rightarrow 0$ and $\mathbf{K}_{ll,a,1} = \mathbf{K}_{ll,a,2} = \ldots \mathbf{K}_{ll,a,v} \rightarrow 0$, with subscripts $(..)_{e,v}$ and $(..)_{a,v}$ indicating the variance-covariance matrices for parameter sets $e$ and $a$, corresponding to view number $v$.

It is to mention here, that using the Bundle Adjustment procedure for pure spatial intersection or for the estimation of camera parameters only, error propagation for the new set of parameters is biased. This is due to the fact that, e.g., with $\mathbf{K}_{ll,a,v} \rightarrow 0$, the components $'a'$ are supposed to be error free observations. Thus when one or the other extreme case is used, the new parameters are preferably estimated in the form of $f(\hat{\mathbf{l}}, \hat{\mathbf{p}}) = \hat{\mathbf{v}}$. More details are given in Appendix C.

### 4.4.2 Significance Test for Additional Parameters

The additional parameters (AP) in the collinearity Eqn. (4.18) have to be tested for significance. As shown in *Gruen* (1986), a null and alternative hypothesis $H_0$ and $H_A$, respectively, are formulated as

$$H_0 : \quad \mathbf{B} \cdot \hat{\mathbf{p}} = 0 \tag{4.21a}$$

$$H_A : \quad \mathbf{B} \cdot \hat{\mathbf{p}} \neq 0 \tag{4.21b}$$

with $\mathbf{B}$ being the coefficient matrix, and $\hat{\mathbf{p}}$ being the corresponding estimated parameter(s) to be tested for significance. If $n$ is the number of additional parameters being tested, the corresponding test quantity is given by:

$$T = \frac{1}{n \cdot \sigma_o^2} \cdot \hat{\mathbf{p}}^T \cdot \mathbf{B}^T \mathbf{Q}_{\hat{p}\hat{p}} \mathbf{B} \cdot \hat{\mathbf{p}} \tag{4.22}$$

The acceptance of $H_0$ is found in the *Fisher*-distribution for a significance level $\alpha$ with the degrees of freedom $n$ and $r$ ($r$ is the redundancy of the system). The alternative test $H_A$ can be used to compute the power of the test $1 - \beta$, i.e., the probability of detecting $H_A$ in case $H_A$ is true. For this purpose, the non-centrality parameter $\delta$, given by

$$\delta = \frac{1}{\sigma_o^2} \cdot \hat{\mathbf{p}}^T \cdot \mathbf{B}^T \mathbf{Q}_{\hat{p}\hat{p}} \mathbf{B} \cdot \hat{\mathbf{p}} \tag{4.23}$$

is computed, where the test follows the non-central *Fister*-Distribution. The higher $\delta$, the better the two tests can be separated, i.e., the power of the test increases. Its maximum is reached, if the set of parameters that are used in the test (4.21) is orthogonal to the remaining parameters involved in the adjustment (e.g. *Gruen*, 1986). As the additional parameters used in the current implementation of the bundle adjustment (Eqn. (4.19)) are given by the expression in (4.15), orthogonality is not given. Thus the strategy used to test the significance of the additional parameters is to first test each additional parameter for its significance[3] (Eqn. (4.22)). If more than one parameter is found to be non-significant, the candidates are tested with respect to each other. Defining a matrix $\mathscr{F}$ that relates each test size $T$ with its candidate parameter $\hat{p}_c$ in the form of

$$\mathbf{V} = \mathscr{F} \cdot \hat{\mathbf{p}}_c \tag{4.24}$$

As each parameter is tested separately, a diagonal form of $\mathscr{F}$ can be found as:

$$\mathscr{F} = \begin{pmatrix} \frac{1}{\sqrt{\mathbf{U}_{11}}} & & & \\ & \frac{1}{\sqrt{\mathbf{U}_{22}}} & & \\ & & \ddots & \\ & & & \frac{1}{\sqrt{\mathbf{U}_{ii}}} \end{pmatrix} \tag{4.25}$$

$$\text{with} \quad \mathbf{U} = \mathbf{B}^T \mathbf{Q}_{\hat{p}\hat{p}} \mathbf{B} \tag{4.26}$$

By solving the eigenvalue problem of the matrix $\mathscr{F} \cdot \mathbf{U} \cdot \mathscr{F}'$, the linear combination of the eigenvector elements corresponding to the largest eigenvalue indicate the relative contribution of each parameter, thus indicating the parameter that has the weakest significance.

### 4.4.3 Gross Error Detection in Image Coordinates

As described in Section 3.2.3, a global test can be carried out to evaluate, if the a posteriori variance, $\hat{\sigma}_0^2$, is equal to the a priori variance, $\sigma_0^2$. If this test fails, gross errors are most likely (also non-Gaussian distribution, systematic errors, wrong a priori covariance matrix, or wrong functional model are other possible explanations). The derivation of gross error detection is found extensively in the literature. For example in *Gruen* (1986) and *Guillaume* (2014), derivations of gross error detection are shown. Here, only the most important results and concepts are given.

For a set of observations $\mathbf{l}$ described by a Gaussian random vector, the following relation holld w.r.t. its true observations:

$$\mathbf{l} \sim \mathcal{N}(\check{\mathbf{l}}; \sigma_0^2 \cdot \mathbf{Q}_{ll}) \tag{4.27}$$

whereas with a gross error (denoted by $\Delta\mathbf{l}$) this becomes:

$$\mathbf{l} \sim \mathcal{N}(\check{\mathbf{l}} + \Delta\mathbf{l}; \sigma_0^2 \cdot \mathbf{Q}_{ll}) \tag{4.28}$$

The consequence of this gross error is that the parameters $\mathbf{p}$, the residuals $\mathbf{v}$, and the a posteriori variance $\hat{\sigma}_0^2$ will be biased. To detect possible gross errors in image coordinate

[3] after the non-linear system has converged

measurements, a two-sided hypothesis test has to be carried out for each observation candidate $i$:

$$H_{0i}: \quad \Delta\hat{l}_i = 0 \tag{4.29a}$$

$$H_{Ai}: \quad \Delta\hat{l}_i \neq 0 \tag{4.29b}$$

The critical test size is then given by

$$T_i = \frac{\Delta\hat{l}_i^2}{\sigma^2_{\Delta\hat{l}_i}} \tag{4.30}$$

which follows the *Fisher*-distribution with degrees of freedom 1 and $r-1$ and, for the two-sided test $T_i^{1/2}$, the *Student*-Distribution, $T_i^{1/2} \sim t(r-1)$. As e.g., shown in *Gruen* (1986), the test size can be computed as:

$$T_i^{1/2} = \frac{\Delta\hat{l}_i}{\sigma_{\Delta\hat{l}_i}} = -\frac{\Delta\hat{v}_i}{\sigma_0 \cdot \sqrt{q_{\hat{v}_i\hat{v}_i}}} = -w_i \tag{4.31}$$

$$\text{using} \quad \Delta\hat{l}_i = -\frac{q_{l_il_i}}{q_{\hat{v}_i\hat{v}_i}} \cdot \Delta\hat{v}_i \tag{4.32}$$

$$\sigma_{\Delta\hat{l}_i} = \hat{\sigma}_0 \cdot \frac{q_{l_il_i}}{\sqrt{q_{\hat{v}_i\hat{v}_i}}} \tag{4.33}$$

where $q_{\hat{v}_i\hat{v}_i}$ is the $i$-th diagonal element of the matrix $\mathbf{Q}_{\hat{v}\hat{v}}$. For a significance level of $\alpha = 1\%$, the critical test size $T_i^{1/2}$ is 2.60.

The principle of (4.31) is to detect a single gross error. *Gruen* (1986) derived a formulation for detecting multiple gross errors simultaneously by adjusting the formulation of the hypothesis $H_0$, i.e.:

$$H_{0i}: \quad \mathbf{B} \cdot \Delta\hat{l} = 0 \tag{4.34}$$

$$H_{Ai}: \quad \mathbf{B} \cdot \Delta\hat{l} \neq 0 \tag{4.35}$$

which leads to a test size of

$$T = \frac{1}{n \cdot \sigma_0^2} \cdot \Delta\hat{\mathbf{v}}^T \mathbf{Q}_{ll}^{-1}\mathbf{B} \cdot \left(\mathbf{B}^T\mathbf{Q}_{ll}^{-1}\mathbf{Q}_{vv}\mathbf{Q}_{ll}^{-1}\mathbf{B}\right)^{-1} \cdot \mathbf{B}^T\mathbf{Q}_{ll}^{-1}\Delta\hat{\mathbf{v}} \tag{4.36}$$

The restriction here is that not all combinations of errors can be tested, as the regularity of $\mathbf{B}^T\mathbf{Q}_{ll}^{-1}\mathbf{Q}_{vv}\mathbf{Q}_{ll}^{-1}\mathbf{B}$ is the limiting condition.

### 4.4.4 Internal/External Reliability

The alternative hypothesis $H_A$ (in Eqn. (4.29)) can be used to assess the internal reliability of the system, i.e., the minimum detectable bias in an observation $l_i$. Defining $\alpha = 1\%$ and $\beta = 5\%$[4], the non-centrality parameter $\lambda$ yields 4.2. With this, the minimum detectable bias in observation $i$, $\nabla l_i$, is given by:

$$\nabla l_i = \lambda \cdot \sigma_{\Delta \hat{l}_i} = \lambda \cdot \sigma_0 \cdot \frac{q_{l_i l_i}}{\sqrt{q_{\hat{v}_i \hat{v}_i}}} = \lambda \cdot \frac{\sigma_{l_i}}{\sqrt{z_i}} \tag{4.37}$$

$$\text{with} \quad z_i = \frac{q_{\hat{v}_i \hat{v}_i}}{q_{\hat{l}_i \hat{l}_i}} \tag{4.38}$$

Relation (4.37) can now be used to estimate the external reliability, i.e., the impact of non-detectable gross errors on the unknown parameters $\hat{\mathbf{p}}$. In the Gauss-Markov model (Eqn. (4.19)), this is given by

$$\nabla \hat{\mathbf{p}}_i = \mathbf{Q}_{\hat{x}\hat{x}} \cdot \mathbf{A}^T \cdot \mathbf{Q}_{ll}^{-1} \cdot \nabla \mathbf{l}_i \tag{4.39}$$

$$\text{with} \quad \nabla \mathbf{l}_i = \begin{pmatrix} 0 & \dots & 0 & \nabla l_i & 0 & \dots & 0 \end{pmatrix}^T \tag{4.40}$$

$\nabla \hat{\mathbf{p}}_i$ is a vector showing the influence of the undetected gross error $\nabla l_i$ on each parameter. The complete influence matrix $\mathcal{N}$ is then obtained by

$$\mathcal{N} = \mathbf{Q}_{\hat{x}\hat{x}} \cdot \mathbf{A}^T \cdot \mathbf{Q}_{ll}^{-1} \cdot \begin{pmatrix} \nabla l_1 & 0 & \dots & 0 \\ 0 & \nabla l_2 & & \vdots \\ \vdots & & \ddots & 0 \\ 0 & \dots & 0 & \nabla l_n \end{pmatrix} \tag{4.41}$$

$\mathcal{N}$ is used to find the maximum influence and its associated observation. This result is helpful for analyzing the reliability of a network (any geodetic, photogrammetric, etc. network) or to optimize an existing network by improving specific observations.

### 4.4.5 Detecting Errors in Control Points

Control points are usually used to define the datum of the network (e.g. *Kraus*, 1997). If there is a sufficient redundancy in the distribution and number of control point coordinates, the estimated coordinates of 'free' control points can be tested against their measured value (typically control points are measured by other geodetic methods - e.g. GNSS (Section 4.5.3), yielding measurements of higher precision. The hypothesis $H_0$ can then be formulated as:

$$H_{0,x}: \quad X_{ph} - X_{geo} = 0 \tag{4.42a}$$

$$H_{0,y}: \quad Y_{ph} - Y_{geo} = 0 \tag{4.42b}$$

$$H_{0,z}: \quad Z_{ph} - Z_{geo} = 0 \tag{4.42c}$$

with $(..)_{ph}$ and $(..)_{geo}$ being the photographically and geodetically determined coordinates, respectively. These one-dimensional test quantities follow the *Student*-Distribution.

[4] these are the traditional levels in geodesy (*Guillaume*, 2014)

### 4.4.6 Model Errors

Ignoring systematic errors in the functional model cause the estimated parameters as well as their variance components to be biased. In principle there are three main types of model error: local, regional and global errors. With respect to image projection, the spatial dimension of the projection error gives the error type:

- Local errors arise, e.g., from local asperities of the imaging sensor surface.
- Regional errors can be present f.e. due to varying atmospheric densities causing complex atmospheric refraction patterns that influence the ray traveling path.
- Global errors result, e.g., due to the Earth curvature (if not taken into account - see Section 4.1), lens distortion, atmospheric refraction.

In the current work, only global errors are considered, as these are generally the largest errors that might cause the functional model to be inaccurate (*Gruen*, 1986). As the curvature of the Earth was already taken into account by working in a topocentric coordinate frame, the two remaining issues are lens distortion and atmospheric refraction, discussed in the following paragraphs.

#### Lens Distortion

As introduced in Section 4.3.2, lens distortion is modeled by a polynomial function of degree 5. *Brown* (1976) derived the form of this polynomial (Eqn. (4.15)) based on the approach to physically model image deformation as good as possible. In contrast, one may also define an arbitrary polynomial that optimally removes systematic lens distortion, if a specific arrangement of measurement points is given. In the latter case, polynomials should be formulated such that optimal orthogonality is achieved (this increases the power of the test when testing the parameter significance). *Gruen* (1986) suggests to use bivariate polynomials with a degree that matches the density and distribution of image points.

For the current situation, intrinsic camera parameters were estimated using Brown's 5 polynomial model (see Section 4.5.3). For the current stage, this is the polynomial used to compensate for lens distortion effects. A transformation and/or re-estimation of a bivariate polynomial may be an interesting approach for future studies.

#### Atmoshperic Refraction

Observations of fixed points at large distances are affected by atmospheric refraction. The exact path of a ray between points $A$ and $B$ is given by Fermant's Principle:

$$\frac{1}{c}\int_A^B ds = \int_A^B dt \rightarrow min \tag{4.43}$$

$$\text{with} \quad c = \frac{c_v}{n} \tag{4.44}$$

with $c_v$ as the speed of light in vacuum and $c$ as the speed of light in a medium with a refractive index $n$. $n$ is a function of wavelength, space, time, pressure, temperature, and air moisture: $n = f(\lambda, \mathbf{X}, t, P, T, e)$. The curve the light path takes can be computed, if the field $n$ is known. For a simple atmospheric layering, $n$ is defined analytically as a function of height and refraction index $n_A$ at point $A$. As shown in *Elmiger* (2002), this leads to a light path following a circle. Depending on the atmospheric layering, the refractive index $k$ (defined as the radius of the light curve with respect to the curvature of the Earth) varies between $k = 1.3$ for near surface layers and $k = 0.155$ for high altitudes. Following *Elmiger* (2002), the influence of the atmospheric refraction on the height component, $\Delta Z$, can be expressed as:

$$\Delta Z = k \cdot \frac{\Delta X^2 + \Delta Y^2}{2R_{Earth}} \tag{4.45}$$

with $\Delta X, Y$ describing the East and North position difference and $R_{Earth}$ representing the radius of the Earth. For $k = 0.13$, the influence on the height component is about 1 *cm* at a distance of 1 *km*. As the refraction index in an alpine environment is likely to be smaller, its influence is well below the accuracy of the object point reconstruction (see Section 4.4.7), and thus neglected in the reconstruction process.

### 4.4.7 Limitations and Expected Accuracy

This section aims at giving some examples of the capabilities of object point reconstruction concerning to precision and reliability using two or more cameras. In combination with the reconstruction process camera parameters are estimated. Although these parameters are usually of secondary interest, an analysis of their determinability is important for scenarios, where parameters have to be included as pseudo-observations.

The setup of the test field to be considered here is constructed by regularly spaced grid points in a 3D volume in front of the cameras (Fig. 4.8). Synthetic image points are generated by projecting object points into each view. Resulting image coordinates were disturbed by additive Gaussian noise $\mathbf{i} \sim \mathcal{N}(0, \sigma_i)$.

#### Camera Parameter Estimation

Camera parameters from all views in all configurations (Fig. 4.8) can be estimated using the set of synthetic object points and the corresponding image points in each view. Without any a priori information on the extrinsic and intrinsic parameters, an initial parameter estimation must be carried out (see Section 4.5). As all object points are treated as observations of very high accuracy, correlations between different views vanish and therefore the problem can be divided into individual sub-problems. Thus only configuration 3 is used for testing.

Table 4.2 shows how the determinability of camera parameters (specially the higher-order intrinsic parameters $k_2$, $k_3$, $p_1$, and $p_2$) becomes non-significant, as the image observation noise increases ($\sigma_i$). The same conclusion holds for weaker geometries[5]. The test is

[5] weak in the context of a bad image coverage (e.g., only points in the lower half of the image)

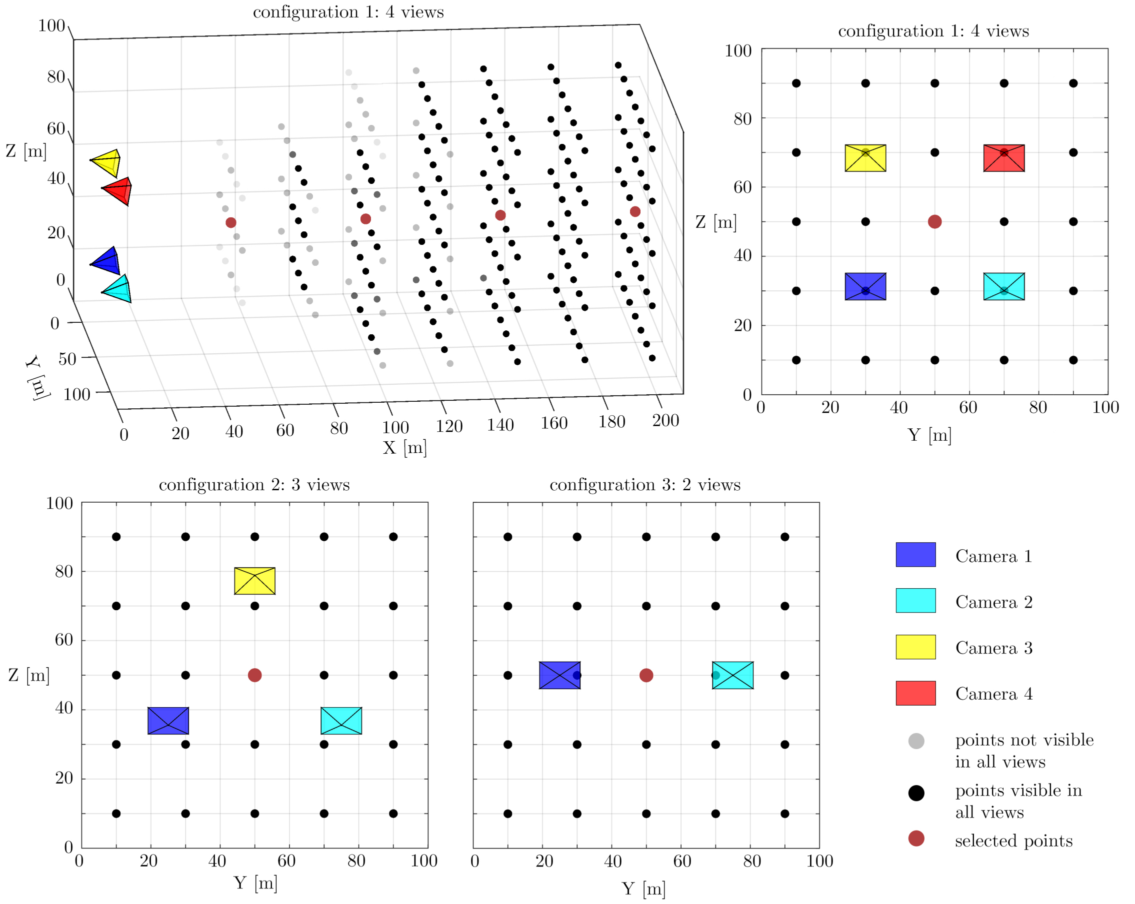


**Figure 4.8:** Setup of the synthetic test summarized in Table 4.2 and 4.3. All baselines between the cameras (except camera 2 to camera 4 and camera 1 to camera 3 in configuration 1) are 50 $m$. Three quadratic backside views show the camera arrangements for the free configurations.

carried out using object point coordinates with the corresponding image point coordinates for each view. For the true camera parameters $\check{p}$, values with the same order of magnitude as determined for the real cameras used in this study, were used (Appendix B). Initial extrinsic parameters are computed by the direct linear transformation method (explained in Section 4.5.1), initial intrinsic parameters were all set to zero. When non-significant parameters are constrained, the standard deviation for the remaining parameters usually increases (see e.g., $p_2$ between noise level $\sigma_i = 0.200$ and $\sigma_i = 0.600$). This also shows the importance of using a robust step-wise parameter elimination scheme to reliably identify non-significant parameters.

### Object Point Reconstruction

Table 4.3 summarizes the results for pure object point reconstruction using the synthetic image points and the given (and fixed) camera parameters. The camera parameters applied here, are listed in Appendix C.3. The estimated standard deviations of the reconstructed

**Table 4.2:** Results of camera parameter reconstruction using synthetic data. For each image noise level $\sigma_i$, all extrinsic, intrinsic, and additional parameters are estimated. The first columns, $\hat{p}$, of each block are the estimated parameters, whereas the second columns, $\hat{\sigma}_p$, show the respective parameter standard deviations. Fields marked by $'-'$ were found not to be significant during the adjustment (see Section 4.4.2), and were thus constrained to zero.

| parameter | $\breve{p}$ | $\sigma_i = 0.200$ | | $\sigma_i = 0.600$ | | $\sigma_i = 1.500$ | |
|---|---|---|---|---|---|---|---|
| | | $\hat{p}$ | $\hat{\sigma}_p$ | $\hat{p}$ | $\hat{\sigma}_p$ | $\hat{p}$ | $\hat{\sigma}_p$ |
| $x_p$ | 0.100 | 0.105 | 0.004 | 0.094 | 0.003 | 0.096 | 0.007 |
| $y_p$ | $-0.100$ | $-0.950$ | 0.003 | $-0.105$ | 0.008 | $-0.063$ | 0.011 |
| $c$ | 15.000 | 15.000 | 0.002 | 15.000 | 0.005 | 14.992 | 0.006 |
| $k_1$ | $1.0e^{-4}$ | $1.06e^{-4}$ | $0.07e^{-4}$ | $0.93e^{-4}$ | $0.00e^{-4}$ | - | - |
| $k_2$ | $-2.0e^{-6}$ | $-2.15e^{-6}$ | $0.16e^{-6}$ | $-1.98e^{-6}$ | $0.48e^{-6}$ | - | - |
| $k_3$ | $-3.5e^{-8}$ | $3.59e^{-8}$ | $0.01e^{-8}$ | $3.52e^{-8}$ | $0.30e^{-8}$ | $2.38e^{-8}$ | $0.06e^{-8}$ |
| $p_1$ | $-1.0e^{-5}$ | $-1.96e^{-5}$ | $0.65e^{-5}$ | - | - | - | - |
| $p_2$ | $8.5e^{-5}$ | $8.14e^{-5}$ | $0.46e^{-5}$ | $7.48e^{-5}$ | $0.01e^{-5}$ | - | - |
| $\omega$ | 1.7000 | 1.7000 | 0.0002 | 1.7000 | 0.0005 | 1.6998 | 0.0007 |
| $\phi$ | 0.0000 | $-0.0003$ | 0.0002 | $-0.0004$ | 0.0002 | $-0.0002$ | 0.0005 |
| $\kappa$ | 0.0000 | $-0.0000$ | 0.0000 | $-0.0000$ | 0.0000 | $-0.0000$ | 0.0000 |
| $O_X$ | 0.000 | 0.009 | 0.007 | $-0.038$ | 0.019 | $-0.041$ | 0.043 |
| $O_Y$ | 50.000 | 50.000 | 0.002 | 50.000 | 0.006 | 50.000 | 0.015 |
| $O_Z$ | 50.000 | 50.010 | 0.002 | 50.022 | 0.007 | 50.001 | 0.015 |
| $\hat{\sigma}_0$ | - | 0.2023 | | 0.6403 | | 1.4983 | |

coordinates match the empirical standard errors between the true and the reconstructed positions within the numerical accuracies. The external reliability indicator shows the sensitivity of the system to undetected gross errors.

With increasing distance, both the standard deviation and the external reliability get worse (i.e., higher values). A difference between 4 and 3 views is found mainly in the external reliability: although standard deviations for the estimated coordinates ($\hat{\sigma}_{X,Y,Z}$) show a minor difference, the external reliability almost doubles for the $X$-component. With 2 views, the standard deviation as well as the external reliability for the $X$-component increase significantly. In this configuration, coordinate components in the line-of-sight direction are weakly bounded and thus image residuals parallel to the baseline (or more generally parallel to the epipolar plane[6]) between the cameras (here it is the image x-axis) become very small (see Fig. 4.9). The consequence is, that the errors of the corresponding coordinate components increase (here the $X$-axis, compare Table 4.3).

[6]The epipolar plane is the plane defined by the too camera centers and the object point seen by both views

**Table 4.3:** Results of object point reconstruction using synthetic data. The 'distance to point' column indicates the distances of the four selected points with respect to the centroid of all cameras. The baseline between the cameras was 50 m in all cases and the image scales for the four distances are: 2.3, 4.2, 6.1, and 8.1 cm/pixel.

| | Distance to Point $[m]$ | Standard Deviation $[cm]$ $\hat{\sigma}_X$ | $\hat{\sigma}_Y$ | $\hat{\sigma}_Z$ | Ext. Reliability $[cm]$ $\nabla X$ | $\nabla Y$ | $\nabla Z$ |
|---|---|---|---|---|---|---|---|
| 4 views | 50 | 1.01 | 0.54 | 0.57 | 1.48 | 1.11 | 1.15 |
| | 100 | 3.81 | 1.05 | 1.07 | 5.54 | 2.14 | 2.18 |
| | 150 | 8.40 | 1.56 | 1.57 | 12.18 | 3.18 | 3.21 |
| | 200 | 14.78 | 2.07 | 2.08 | 21.38 | 4.22 | 4.23 |
| 3 views | 50 | 1.09 | 0.60 | 0.62 | 2.56 | 1.49 | 1.56 |
| | 100 | 4.12 | 1.16 | 1.18 | 9.93 | 2.87 | 2.93 |
| | 150 | 9.09 | 1.73 | 1.74 | 22.12 | 4.26 | 4.30 |
| | 200 | 16.02 | 2.29 | 2.30 | 39.06 | 5.64 | 5.67 |
| 2 views | 50 | 1.45 | 0.72 | 0.72 | 4.20 | 2.10 | 2.10 |
| | 100 | 5.80 | 1.45 | 1.45 | 16.82 | 4.20 | 4.20 |
| | 150 | 13.04 | 2.17 | 2.17 | 37.84 | 6.30 | 6.30 |
| | 200 | 23.14 | 2.89 | 2.89 | 67.16 | 8.40 | 8.40 |

#### Combined Camera and Object Point Estimation

The task of the bundle adjustment is to combine the previously separated estimation procedures. In the given simulation setup, configuration 3 was used to test its performance: 15 control points were randomly selected from the synthetic object points. Together with the image point measurements ($\sigma_i \sim \mathcal{N}(0, \sigma_i = 0.5\,\text{pixel})$), initial camera parameters and object point coordinates were computed. For the same evaluation points at distances of 50, 100, 150, and $200\,m$, equal estimates of parameter precision were obtained, whereas the external reliability is $\sim 15\,\%$ worse than in the case of pure spatial intersection. This result reflects the fact that if there were undetected gross errors in image coordinates, they also affect camera parameters and lead, therefore, to a higher impact on the object point reconstruction. For the camera parameters, the same conclusion holds.

## 4.5 Initial Parameters for Bundle Adjustment

The previous section (Section 4.3.2) showed that the mathematical model used to solve the bundle adjustment is highly non-linear. As such, initial estimates for the internal and external camera parameters are needed. If no estimates for the additional parameters are available (e.g., using a separate calibration procedure), they are set to zero as this has only a secondary effect (i.e., for the reduction of systematic errors (*Gruen*, 1986)). Along with the intrinsic parameters, the extrinsic parameters (camera position and orientation) have to be estimated for each view in the first place. Afterwards, the initial object coordinates can be estimated.

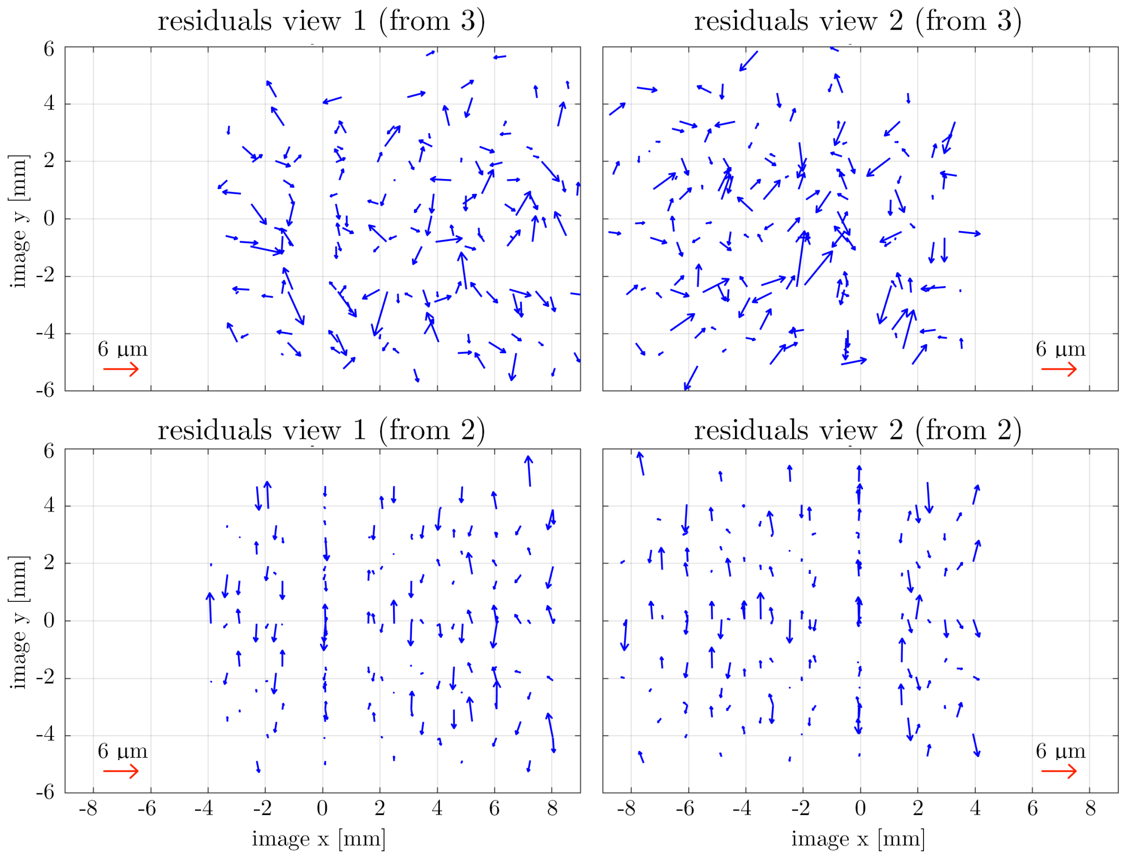


**Figure 4.9:** Image residuals for configuration 2 (3 views - top figures) and configuration 3 (2 views - bottom figures). For configuration 2, only the first two image residuals are shown. The vector scale of 6 $\mu m$ corresponds to 1 pixel. $\sigma_i = 0.5$ pixel in both cases. Only residuals of points seen from all views are plotted.

### 4.5.1 Intrinsic/Extrinsic Parameters

In the general case of a projective camera, a matrix $\mathbf{P}$ maps the object points (3D space) $\mathbf{X} = [X\ Y\ Z\ 1]^T$ to the corresponding image points (2D space) $\mathbf{x} = [x\ y\ 1]$:

$$\mathbf{x} = \mathbf{P} \cdot \mathbf{X} \tag{4.46}$$

Matrix $\mathbf{P}$ thus encodes all the mapping properties of the camera. Following *Hartley and Zisserman* (2003), the $4 \times 3$ $\mathbf{P}$-matrix is defined as:

$$\mathbf{P} = \begin{pmatrix} p_{11} & p_{12} & p_{13} & p_{14} \\ p_{21} & p_{22} & p_{23} & p_{24} \\ p_{31} & p_{32} & p_{33} & p_{34} \end{pmatrix} \tag{4.47}$$

Eqn. (4.46) above can also be written in the form of the cross-product $((\mathbf{x} \times \mathbf{P}) \cdot \mathbf{X} = \mathbf{0})$ such that a simple linear solution can be found:

$$\begin{pmatrix} \mathbf{0}^T & -z_i\mathbf{X}_i^T & y_i\mathbf{X}_i^T \\ z_i\mathbf{X}_i^T & \mathbf{0}^T & -x_i\mathbf{X}_i^T \end{pmatrix} \cdot \begin{pmatrix} p_{11} \\ p_{12} \\ p_{13} \\ p_{14} \\ p_{21} \\ \vdots \\ p_{34} \end{pmatrix} = \mathbf{A} \cdot \mathbf{p} = \mathbf{0} \tag{4.48}$$

with $(x_i, y_i, z_i)$ being the $i$-th homogeneous image point coordinates and $\mathbf{X}_i$ the coordinates of the corresponding object point. Note that following the cross-product derivation, the first matrix in Eqn. 4.48 has an additional row. Due to the linear dependency of this row with respect to the first rows, it can be dropped (see *Hartley and Zisserman* (2003) for more details). Using at least six control points ($i = 1..6$), the linear equation can be solved for all the elements of $\mathbf{P}$. With $n$ observation pairs $\mathbf{x}_i \leftrightarrow \mathbf{X}_i$, $i = 1..n$, the system matrix has dimension $2n \times 12$. As matrix $\mathbf{P}$ has 11 degrees of freedom[7], an additional condition has to be defined. In a first step, the algebraic error is minimized with the constraint $\|\mathbf{p}\| = 1$. This type of estimation is known as the Direct Linear Transformation (DLT). The solution is the eigenvector of $\mathbf{A}^T\mathbf{A}$ having the smallest eigenvalue (least-squares solution of homogeneous equations). An elegant way to obtain this is to perform a singular value decomposition, where the solution vector is the eigenvector corresponding to the smallest eigenvalue.

As the minimization of the algebraic error does not minimize the geometric distance, such a (nonlinear) minimization can be performed additionally. Using the result obtained with the DLT, a least-squares adjustment minimizing the geometric error can be performed.

$$\sum_i d(\mathbf{x}_i, \mathbf{P} \cdot \mathbf{X}_i)^2 \rightarrow min \tag{4.49}$$

In addition, constraints with respect to known parameters can be included. Typically the camera pixel sizes in $x$- and $y$-direction ($\alpha_x$, $\alpha_y$) are equal and the skew $s$ (see Eqn. 4.52) is constrained to zero. As small variations in these parameters can have a large impact on other parameters, the solution is obtained by iteratively increasing the weight $w$ on the pseudo-observations defining the constraints:

$$\sum_i d(\mathbf{x}_i, \mathbf{P} \cdot \mathbf{X}_i)^2 + w \cdot (\alpha_x - \alpha_y)^2 + w \cdot s^2 \rightarrow min \tag{4.50}$$

In the implementation used in this work, the second adjustment step (non-linear least-squares) is performed by adjusting the parameters of the decomposed camera matrix $\mathbf{P}$. The decomposed elements are obtained by solving for the following relationship:

$$\mathbf{P} = \mathbf{KR}[\mathbf{I} \mid -\mathbf{O}] \tag{4.51}$$

[7] 3 rotation components, 3 translation components, focal lengths $f_x$ and $f_y$, skew $s$, and principle point offset $x_p$, and $y_p$

with

$$\mathbf{K} = \begin{pmatrix} f_x & s & x_0 \\ 0 & f_y & y_0 \\ 0 & 0 & 1 \end{pmatrix} \qquad \text{calibration matrix} \tag{4.52}$$

$$\mathbf{R} = \begin{pmatrix} r_{11} & \cdots & r_{13} \\ \vdots & \ddots & \\ r_{31} & & r_{33} \end{pmatrix} \qquad \text{rotation matrix} \tag{4.53}$$

$$\mathbf{O} = \begin{pmatrix} O_X \\ O_Y \\ O_Z \\ 1 \end{pmatrix} \qquad \text{camera center} \tag{4.54}$$

$\mathbf{I}$ is the $3 \times 3$ identity matrix, $f_x$ and $f_y$ are the respective focal lengths of the camera, $s$ is the skew parameter, and $x_o$ and $y_o$ represent the principle point coordinates. If pixel coordinates for image point measurements are used, then the unit of these parameters is pixel. The units of the camera center (in the world coordinate system), $\mathbf{O}$, correspond to the units in $\mathbf{X}_i$. Attention must be paid to use consistent coordinate systems, i.e., either both, the world and the camera coordinates have to be defined in a right- or left-handed coordinate system (Section 4.1). Metric image coordinates can be used in case the pixel size and image dimensions are known. The resulting calibration matrix $\mathbf{K}$ then directly indicates the focal length(s) ($f_x$, $f_y$) and the principle point offset ($x_0$, $x_0$) in *mm* as used for the initial estimates. In this case, however, the direction of the $y$-axis is flipped (points upward). To account for this change in the orientation of the coordinate system, the second column of $\mathbf{K}$ and the second row of $\mathbf{R}$ have to change sign.

The decomposition of $\mathbf{P}$ is straight forward: defining $\mathbf{M} = \mathbf{KR}$ and thus $\mathbf{P} = [\mathbf{M} \mid -\mathbf{MO}]$ (Eqn. (4.51)), it follows that:

$$\mathbf{O} = -\mathbf{M}^{-1} \cdot \mathbf{p}_4 \tag{4.55}$$

with $\mathbf{p}_4$ being the last column of $\mathbf{P}$. $\mathbf{M}$ then has to be further decomposed into an upper triangular matrix $\mathbf{K}$ and an orthogonal matrix $\mathbf{R}$, known as the QR-decomposition. The inversion of $\mathbf{M}$ is only possible, if $\mathbf{M}$ is not singular, that is for a camera at a finite distance ($\mathbf{M}$ is singular for cameras at infinity). One last decomposition - if needed - can be made on $\mathbf{R}$ to retrive the three Euler angles $\omega$, $\phi$, and $\kappa$ (see Equn. (4.1)). Here an adjusted decomposition of the procedure presented in *Slabaugh* (1999) is used.

Table C.2 shows the difference between the solution of the DLT and the constrained least-squares minimization (CLS), (Eqn. (4.50)). Several parameters show differences between the two approaches. The coordinate $O_X$ for camera 2 shows a shift of almost 3 meters and the Euler angle $\omega$ changes by about 17 degrees. Using the CLS as a second step to estimate initial parameters might thus show a considerable improvement.[8] In terms of parameter accuracy, the standard deviations for the CLS adjustment have been computed

[8] For a more detailed analysis of this effect, tests using different initial estimates for the Bundle Adjustment should be performed.

**Table 4.4:** Results of initial estimates on basic camera parameters using 11 ground control points. The DLT (Direct Linear Transformation) columns show the results obtained by minimizing the algebraic error, whereas the CLS (Constrained Least-Squares) columns show the estimated parameters after the non-linear adjustment with $f_y = f_x$ and $s = 0$ imposed as soft constraints (see text for more details). The estimated parameter standard deviations for the CLS are listed in Table 4.5. A 14 mm lens attached to a Nikon D300s was used.

| Parameters | *DLT* | *CLS* | *DLT* | *CLS* |
|---|---|---|---|---|
| | camera 1 | | camera 2 | |
| $f_x$ $[mm]$ | 13.71 | 13.95 | 13.88 | 14.31 |
| $f_y$ $[mm]$ | 13.34 | 13.95 | 13.61 | 14.31 |
| $s$ $[mm]$ | -0.12 | 0.00 | -0.09 | 0.00 |
| $x_0$ $[mm]$ | 0.015 | 0.01 | 0.04 | -0.065 |
| $y_0$ $[mm]$ | -0.08 | 0.57 | 0.05 | 1.43 |
| $\omega$ $[rad]$ | 4.66 | 4.74 | 4.55 | 4.88 |
| $\phi$ $[rad]$ | 4.16 | 4.17 | 4.39 | 4.39 |
| $\kappa$ $[rad]$ | 3.08 | 3.16 | 2.99 | 3.30 |
| $O_X$ $[m]$ | 0.52 | -1.48 | -32.66 | -35.31 |
| $O_Y$ $[m]$ | -0.52 | -1.10 | 20.25 | 19.59 |
| $O_Z$ $[m]$ | 0.57 | -0.11 | -16.46 | -16.97 |

**Table 4.5:** Estimated standard deviations of parameters determined by CLS (see Table C.2).

| parameter $\hat{\sigma}$ | CLS camera 1 | CLS camera 2 |
|---|---|---|
| $f_x$ $[mm]$ | 0.110 | 0.099 |
| $f_y$ $[mm]$ | 0.000 | 0.000 |
| $s$ $[mm]$ | 0.000 | 0.000 |
| $x_0$ $[mm]$ | 0.089 | 0.126 |
| $y_0$ $[mm]$ | 0.440 | 0.546 |
| $\omega$ $[rad]$ | 0.031 | 0.037 |
| $\phi$ $[rad]$ | 0.002 | 0.001 |
| $\kappa$ $[rad]$ | 0.006 | 0.009 |
| $O_X$ $[m]$ | 0.826 | 0.745 |
| $O_Y$ $[m]$ | 0.532 | 0.309 |
| $O_Z$ $[m]$ | 0.240 | 0.140 |

(see Table 4.5). As the corresponding parameter standard deviations for the DLT are not directly accessible, only the a posteriori variance is compared: for the DLT estimates, these are $1.2 \times 10^{-4}$ mm in average, whereas the CLS adjustment yields an average $\hat{\sigma}_0$ of $1.5 \times 10^{-4}$ mm.

### 4.5.2 Object Point Coordinates

The second set of initial parameters that need to be determined are the coordinates of the (3D) object points. As shown in Eqn. (4.46), the relationship between object points $\mathbf{X}$ and image points $\mathbf{x}$ is described by the projective camera matrix $\mathbf{P}$. Once $\mathbf{P}$ has been determined (see the paragraph above), measured image coordinates of the same object point in at least two views can be used to reconstruct the coordinates of the object point. This problem is known as triangulation (*Hartley and Sturm*, 1997). In principle the same algorithm as in the previous section can be used for this task. For each view, there exists a linear relationship for the object point $i$ in the form of:

$$\begin{pmatrix} x \cdot \mathbf{p}_3^T - \mathbf{p}_1^T \\ y \cdot \mathbf{p}_3^T - \mathbf{p}_2^T \end{pmatrix} \cdot \begin{pmatrix} X \\ Y \\ Z \\ 1 \end{pmatrix}_i = \mathbf{A} \cdot \mathbf{X}_i = \mathbf{0} \tag{4.56}$$

Again the third row can be dropped as this is a linear combination of the first two rows (compare with Eqn. (4.48)). For multiple views $j$, matrix $\mathbf{A}$ becomes:

$$\mathbf{A} = \begin{pmatrix} {}^1x \cdot {}^1\mathbf{p}_3^T - {}^1\mathbf{p}_1^T \\ {}^1y \cdot {}^1\mathbf{p}_3^T - {}^1\mathbf{p}_2^T \\ {}^2x \cdot {}^2\mathbf{p}_3^T - {}^2\mathbf{p}_1^T \\ {}^2y \cdot {}^2\mathbf{p}_3^T - {}^2\mathbf{p}_2^T \\ \vdots \\ {}^jx \cdot {}^j\mathbf{p}_3^T - {}^j\mathbf{p}_1^T \\ {}^jy \cdot {}^j\mathbf{p}_3^T - {}^j\mathbf{p}_2^T \end{pmatrix} \tag{4.57}$$

And the linear system can be solved for $\mathbf{X}_i$ using the singular value decomposition. $\mathbf{p}_n^T$ represents the $n$-th row of camera matrix $\mathbf{P}$. As before, the algebraic error is minimized. *Hartley and Schaffalitzky* (2004) for example note that this procedure is not an optimal solution in terms of minimizing the geometric errors, specially in a scenario of nearly parallel rays, may thus lead to initial guess estimates that are far off the actual solution.

As all measurements are noisy by nature, rays in space do not perfectly intersect. Similar to Eqn. (4.49), a cost function that minimizes the projection error can be defined:

$$\sum_i d(\mathbf{x}_i, \mathbf{P}_i \cdot \mathbf{X})^2 \rightarrow min \tag{4.58}$$

For two views, a popular method was proposed by *Hartley and Sturm* (1997): here the roots of a sixth-degree polynomial function need to be found, where the optimal solution is then determined by an evaluation procedure of the three local minima.

In this work, initial object point coordinates are determined using the singular-value-decomposition in Eqn. (4.57). This solution is then geometrically optimized using a non-linear adjustment (details are given in Appendix C). The underlying principle is identical to its sub-problem in the bundle adjustment. Optimizing single object coordinates before the bundle procedure allows for faster convergence and generally results in a better stability (as each initial optimization is uncorrelated with the other parameters).

### 4.5.3 Additional Parameters

Camera calibration is usually performed in a special envirnoment or as a self-calibration task during a full bundle adjustment (e.g. *Clarke and Fryer*, 1998). As presented in Eqn. (4.15), the additional parameters are polynomials, where the coefficients need to be reliably determined. Often a network geometry used to reconstruct scene objects is different from that required for a comprehensive camera calibration (*Remondino and Fraser*, 2006). As indicated in Section 4.4.7, the network for a reliable self-calibration with additional parameters needs to be very strong, i.e., measurement points in the image have to have a dense and homogeneous distribution and image coordinates need to be measured with sufficient precision (also noted in *Luhmann et al.* (2006)).

Nowadays, camera calibration is an easy task and can be performed in almost all environmental conditions. A self-calibrating bundle adjustment in a local coordinate system is typically carried (*Hartley*, 1994). Here, the camera observes a static scene from various positions. Characteristic markers or point patterns that are commonly used and detected in each image are used to simplify the process of point identification. The next section contains more details about this calibration procedure.

### 4.5.4 Camera Calibration

#### Intrinsic and Additional Parameters

For a reliable estimation of the additional parameters, an in-field calibration was carried out. A calibration board shown in Fig. 4.10 was used after the camera installation in the weatherproof housing box was completed. About 50 images were taken, where the board orientation was changed after every image[9]. This has the same effect as moving the camera around a static scene. The computation of the self-calibrating bundle adjustment in a local coordinate system was done in Australis *Photometrix* (2013). The patterns fixed onto the rigid foam board are identified by their unique point layout. The result obtained for the intrinsic and additional parameters is given in Appendix B.

#### Extrinsic parameters

Absolute orientation, i.e., orientation of the camera with respect to the area of interest was performed after the calibration for the intrinsic and additional parameters was completed. Focus and camera position with respect to the cover glass remained unchanged for this task. Nine ground control points (GCPs) - rectangular orange patterns - were placed in the field and oriented such that they could easily be detected in both views. Each position was measured by GNSS to an average accuracy of $\sigma_x = \sigma_y = 1\,\mathrm{cm}$ and $\sigma_z = 3\,\mathrm{cm}$. Together with the permanent GNSS stations in the field, a total of 11 stations for absolute orientation could be used.

[9] With optimal calibration images, only about 10 images would be necessary to reliably estimate all parameters.

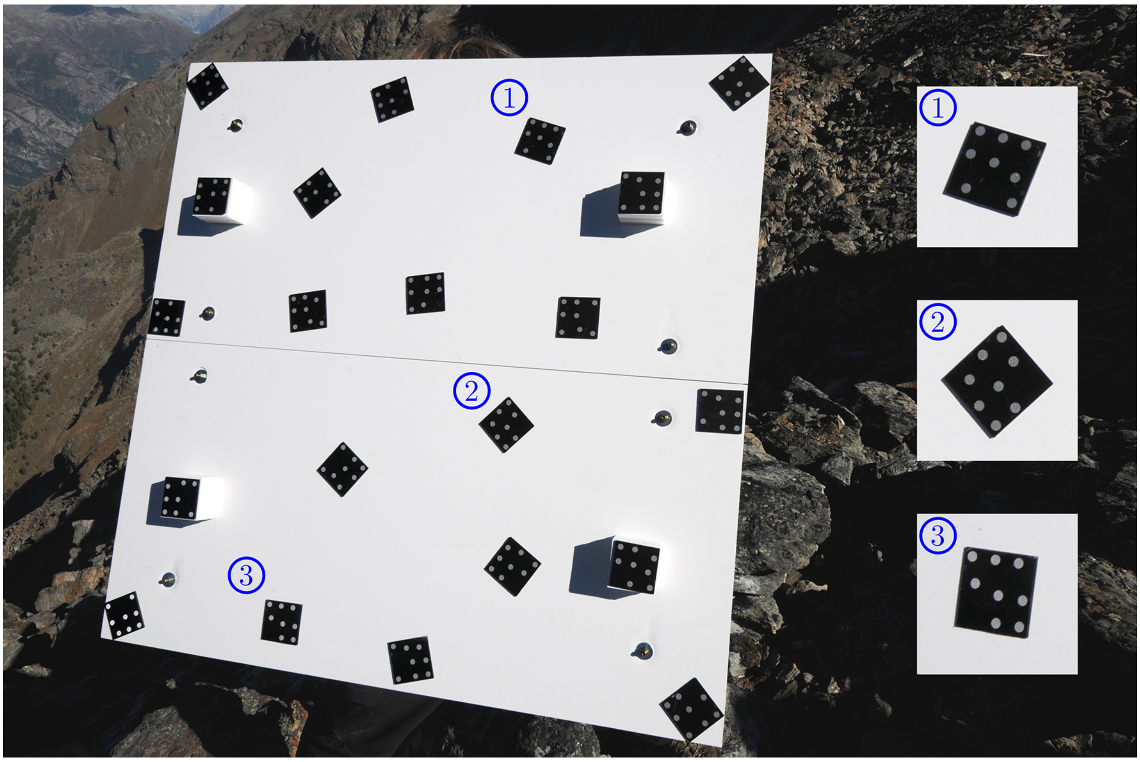


**Figure 4.10:** Calibration board used to estimate intrinsic and additional parameters. Three close-up views of uniquely defined calibration patterns are shown (Nr. 1-3). The estimation procedure was carried out with the Australis software package (*Photometrix*, 2013).

Image coordinates corresponding to the individual stations were determined in a semi-automatic process: initialized by defining approximate coordinates (±100 pixel), a clustering of pixel values based on *kmeans* (*Arthur and Vassilvitskii*, 2007) was used for image segmentation. To increase the efficiency, the color image was converted into gray scales such that the orange calibration plates were easily detected (see Fig. 4.11). For the identification of the calibration pattern, different statistical properties of the segmentation clusters were extracted:

- cluster roundness: a maximum deviation of ±30% from a circle is allowed.
- minimum/maximum size: the calibration pattern size can optionally be constrained to be in a given range, e.g., $5^2 \leq Area \leq 50^2$.
- pixel intensity: the mean luminosity is expected to be in the upper end of the histogram.

For calibration patterns at larger distances small errors in the segmentation process caused the centroids to be inaccurate. To overcome this limitation a region growing was applied at the initial centroid position. An example of such a refinement process is given in Fig. 4.12.

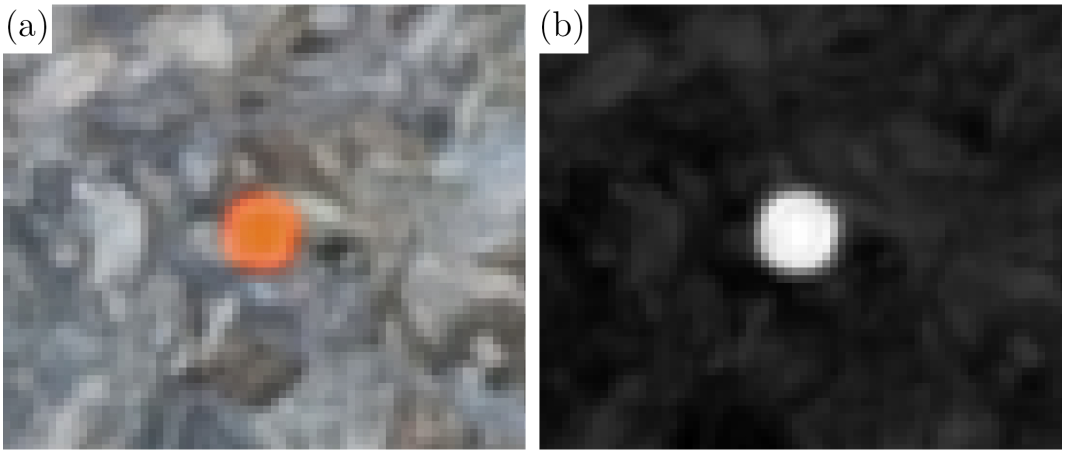


**Figure 4.11:** Example of a calibration plate ($40 \times 40\ cm$) seen in camera 2. (a) shows a crop of the original image, whereas the black-and-white image in (b) was created by $I_{bw} = I_r - 0.4 \cdot I_g - 0.4 \cdot I_b$, with subscripts $r$, $g$, and $b$ indicating the corresponding color channel.

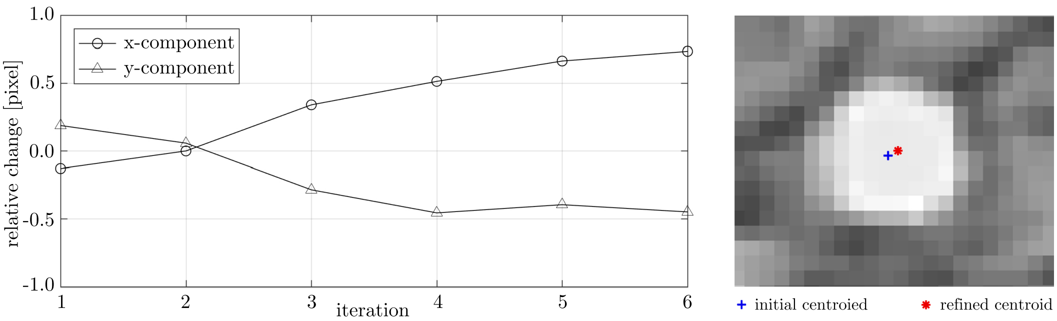


**Figure 4.12:** Refinement of centroid position based on region growth. The left plot shows the change of the centroid after each regional growth for the $x-$ and $y-$components. On the right a highly zoomed image shows the initial and final centroid.

Fig. 4.13 shows the result of the centroid centers determined after the refinement. Visual inspection suggests that the achieved precision is in the order of 0.2-0.5 pixel[10].

As demonstrated in the previous chapter, the precision of the LSM technique is in the order of 0.05 pixel for ideal targets. Thus a template matching procedure could be used instead of the region growing process. The image coordinates corresponding to the respective (known) object coordinates are finally used to perform an initial estimate of the extrinsic parameters (Section 4.5.1).

[10]This value is confirmed by the initial bundle adjustment (Section 6.2.1) with one centroid position being found to be erroneous.

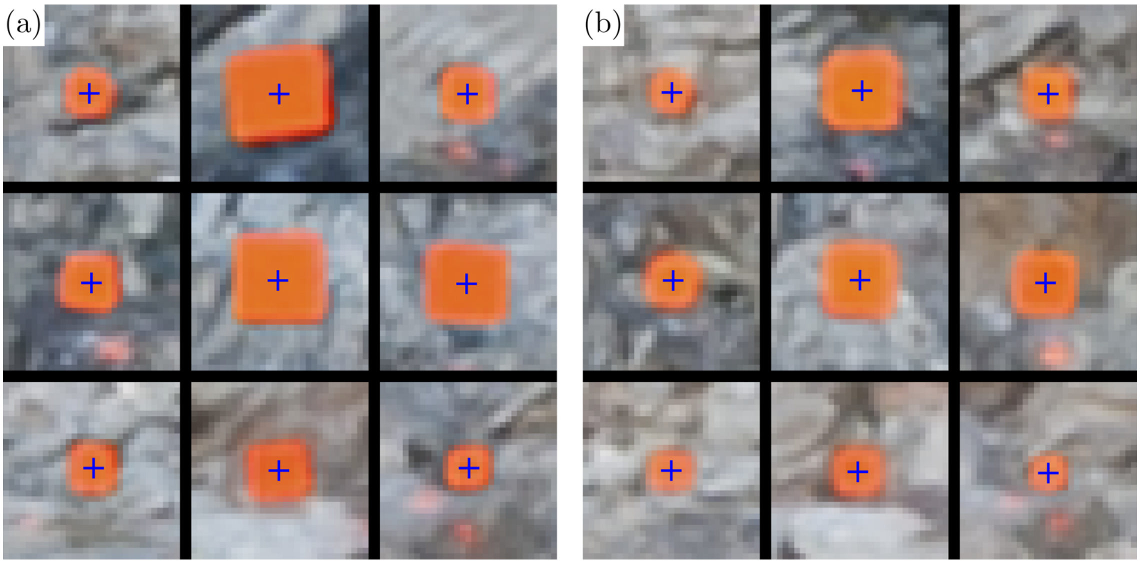


**Figure 4.13:** Centroids of calibration patterns determined in the images. Image coordinates are used to relate the pixel to the world coordinate system.

## 4.6 Integration of Permanent GNSS Stations

As explained in Chapter 1, the goal of this study is to estimate surface motion of a creeping permafrost area over a period of multiple years. As the stability of the camera platform over such a long period was uncertain, each camera station was equipped with a GNSS module to accurately track its motion. In addition, two permanent GNSS stations were installed on large rock boulders of the moving rock slide in the field of view of the cameras (see Section 2.2 for hardware details). The primary goal of these stations is to deliver daily solutions of their respective antenna positions. For the purpose of photogrammetric object point reconstruction they can also be used as accurate control point measurements, if the GNSS antennas can be identified within the images.

An accurate localization of the GNSS antennas in all images is needed to link their actual position to image coordinates. This task is realized by an image template matching procedure, following the principles of LSM described in Section 3.2. Due to the flexibility of the mathematical model, templates of arbitrary shape can be defined (Fig. 4.14). The two GNSS stations in view (called station 1 and station 2) are at mean distances of 128 and 101 m from the two cameras, respectively. For station 2, the outline of the template had to be chosen such that only a minimum number of background pixels are included. This was necessary as the background seen in this projection is at a much larger distance and moves considerably less than the station in the foreground. For the station 1, this was not of concern.

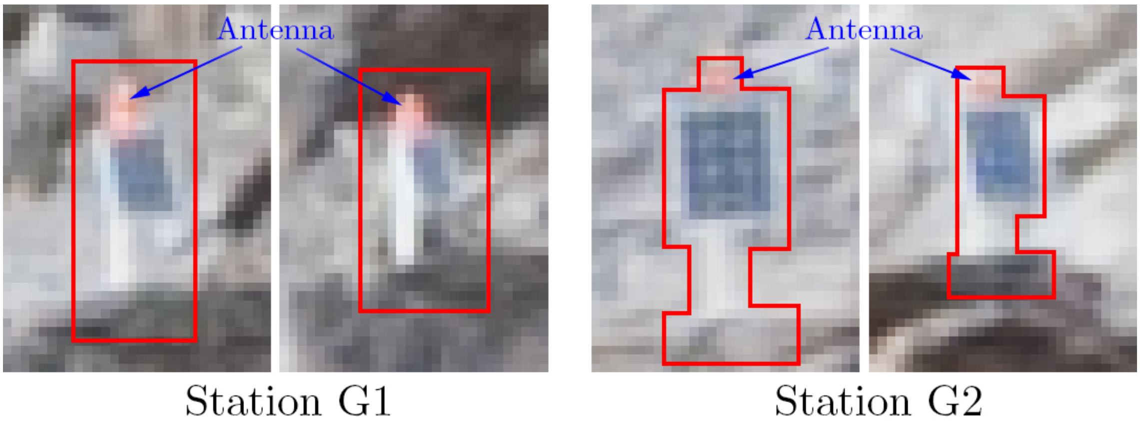


**Figure 4.14:** Low cost GNSS stations seen in the two views. For station 1, the image scales are 5.2 and 6.4 cm/pixel and for station 2, the image scales are 4.1 and 4.5 cm/pixel for camera 1 and 2, respectively (based on the distances from the GNSS antenna to the respective camera centers). Outlined in red are the corresponding templates used for matching. Blue arrows point to the antenna centers.

The transformation being estimated is formulated with respect to the GNSS antenna centers (as these correspond to the measured GNSS coordinates). For a given template image, the pixel coordinates of the corresponding antenna positions were manually determined. To increase redundancy and reduce the error of manual pixel localization, a template was defined for each operation year, yielding multiple matching results for each image. Naturally, a systematic bias can not be excluded at this stage, because (1) the antennas in the images are only about 2 pixel in size and (2) the exact antenna center is not known. Combined, a systematic error in the order of 0.5 pixel or 5 cm (projected) might be present.

As mentioned before, the matching principle is identical to the least-squares technique used for feature tracking. A notable difference is the matching window geometry: within the template outline, the number of pixels near the antennas are very few and thus other areas have to be included to reach a certain level of redundancy and contrast. As the border of the solar panels show the highest image gradients, the matching results are most sensitive to these structures.

## 4.7 Processing Strategy

By combining GNSS and image-based point measurements, the following strategy for object point reconstruction is applied:

- Initial camera calibration, epoch $t_0$: independent estimation of extrinsic, intrinsic, and additional parameters (Section 4.5.4).
- Optimization of all camera parameters for epoch $t_0$: the combined set of parameters are re-estimated using the principle of bundle adjustment (Section 4.4). Parameters that are weakly defined in this geometry are given appropriate weights.
- Estimation of new object points on solid rock for epoch $t_0$: new 3D points are defined to be used as pseudo ground control points for epochs $t > t_0$.
- Estimation of coordinates for features on the surface of the permafrost creep for epoch $t_0$.
- Re-estimation of camera parameters and coordinates of moving features: the principle of bundle adjustment is applied for epochs $t > t_0$. The camera positions are kept at constant offsets with respect to the estimated positions of the onboard GNSS antenna centers. Daily GNSS positioning solutions are interpolated for the respective epochs of image acquisition. The initial set of pseudo ground control points along with the high precision GNSS ground control points are used for every epoch possible.

As noted in Chapter 2, images are typically acquired once every hour. Due to the slow motion of the rock slide, however, only one image per day was selected (manually) for further analysis. Both cameras were given the same image acquisition schedule. Although the clocks are accurate, absolute differences of up to 10 seconds in the image acquisition epochs between the cameras were observed (acquisition time stamp in the image headers). Due to the slow motion of the rock slide, this is not of concern for the reconstruction process. For the automatic allocation of images corresponding to the same epoch, a respective acquisition time tolerance of a few seconds was therefore used.

# 5 Collocation for Time Series Analysis

## 5.1 Least-Squares Collocation

### 5.1.1 Principle

Least-squares collocation is a well known method to differentiate between measurement noise and signal, based on assigned neighbourhood relations (i.e., correlations). In the geodetic content, this method is well described in *Moritz* (1970) and *Moritz* (1973). Traditionally a linear estimator of a general vector field $l(\mathbf{r})$ is combined with an empirical estimate of correlations in a stochastic field $\mathbf{C_{ss}}$. In case the observation noise $\mathbf{n}$ and the statistical signal $\mathbf{s}$ are uncorrelated and $\mathbf{s} \sim \mathcal{N}(0; \mathbf{C_{ss}})$ and $\mathbf{n} \sim \mathcal{N}(0; \mathbf{C_{nn}})$ the solution can be formulated as:

$$\mathbf{l}(\mathbf{r}) = \mathbf{A} \cdot \hat{\mathbf{x}} + \hat{\mathbf{s}} + \hat{\mathbf{n}} \tag{5.1a}$$

$$\mathbf{D} = (\mathbf{C_{ss}} + \mathbf{C_{nn}})^{-1} \tag{5.1b}$$

$$\hat{\mathbf{k}} = \mathbf{D} \cdot (\mathbf{l}(\mathbf{r}) - \mathbf{A} \cdot \hat{\mathbf{x}}) \tag{5.1c}$$

$$\hat{\mathbf{s}} = \mathbf{C_{ss}} \cdot \hat{\mathbf{k}} \tag{5.1d}$$

$$\hat{\mathbf{n}} = \mathbf{C_{nn}} \cdot \hat{\mathbf{k}} \tag{5.1e}$$

where $\mathbf{C_{ss}}$ is the covariance matrix of $\hat{\mathbf{s}}$, $\mathbf{C_{nn}}$ is the (auto-) covariance matrix of $\hat{\mathbf{n}}$, and $\hat{\mathbf{x}}$ the estimated parameter vector of the deterministic part, given in Eqn. (5.2), with $\mathbf{A}$ as its design matrix and $\mathbf{Q}_{xx}$ as the parameter variance-covariance matrix. Matrix $\mathbf{D}$ represents the weight matrix, and $\hat{\mathbf{k}}$ the correlation vector. The system minimizes $\mathbf{n^T C_{nn}^{-1} n} + \mathbf{s^T C_{ss}^{-1} s}$. Also note that vector $\mathbf{r}$ represents coordinates in space and time.

$$\hat{\mathbf{x}} = \mathbf{Q}_{\hat{x}\hat{x}} \cdot \mathbf{A}^{-1} \cdot \mathbf{D} \cdot \mathbf{l}(\mathbf{r}) \tag{5.2}$$

$$\mathbf{Q}_{\hat{x}\hat{x}} = (\mathbf{A^T} \cdot \mathbf{D} \cdot \mathbf{A})^{-1} \tag{5.3}$$

The general (noise free) interpolated field $\mathbf{l}(\mathbf{r}')$ at the predicted coordinates $\mathbf{r}'$ is then given by

$$\hat{\mathbf{l}}(\mathbf{r}') = \mathbf{A}' \cdot \hat{\mathbf{x}} + \mathbf{C}_{s's} \cdot \hat{\mathbf{k}} \tag{5.4}$$

with $\mathbf{C_{s's}}$ as the covariance matrix between the predicted and the measured points. The correlations between the measured positions described in the covariance matrix $\mathbf{C_{ss}}$, the (auto-)correlations of the noise components described in $\mathbf{C_{nn}}$, and their relative weights determine how the measurements are distributed into a signal $\hat{\mathbf{s}}$ and a noise $\hat{\mathbf{n}}$ part. The signal correlation follows an a priori model, either determined by external information (e.g., a known tectonic setting), see *Egli et al.* (2007) or *Villiger* (2014) for further details, or an estimated correlation function computed by using the covariogram representation (*Journel*, 1989). If both, a physical correlation between measurements at different coordinates and a database with units of geometrically well distributed measurement positions at sufficient measurement quality is given, the variance and co-variance components can be determined empirically, see Section 5.2.

### 5.1.2 Derived Quantities

The formulation given by Eqn. (5.2) can also be formulated such that a quantity derived from the observation vector $\mathbf{l}$ is estimated. A typical example, that is also used in this study, is the estimation of the velocity from a sequence of observed positions coordinates. The relation is formulated as the derivative w.r.t. time $\mathbf{t}$:

$$\mathbf{v}(\mathbf{r}) = \frac{\partial \mathbf{l}(\mathbf{t}, \mathbf{r})}{\partial \mathbf{t}} \tag{5.5}$$

For the collocated velocity of the interpolation field $\mathbf{v}(\mathbf{r})$, this is (e.g., *Peter*, 2000):

$$\mathbf{v}(\mathbf{r}) = \nabla \cdot \mathbf{A}' \cdot \hat{\mathbf{x}} + \nabla \mathbf{C}_{s's} \cdot \hat{\mathbf{k}} \tag{5.6}$$

Eqn. (5.6) shows that the derivatives of the deterministic part of matrix $\mathbf{A}$ as well as the derivative of the covariance function $f$ (see Section 5.2), defining the covariance matrix $\mathbf{C}_{s's}$, have to be found.

## 5.2 Empiric Covariance Function

The covariance matrix $\mathbf{C}_{ss}$ is usually described by a correlation function $f$:

$$\mathbf{C}_{ss} = \sigma_s^2 \cdot f(\mathbf{r}_{ij}, r_s) \tag{5.7}$$

with $\mathbf{r}_{ij} = |\mathbf{r}_i - \mathbf{r}_j|$ describing the measurement distance matrix (space and time), $r_s$ the signal correlation length, and $\sigma_s^2$ the signal variance. The solution of Eqn. (5.1) thus highly depends on the choice of the analytical covariance function (5.7), describing the correlation between the measurements. The choice of the function $f$ is not free because it must follow a series of important properties (e.g., *Geiger*, 1996; *Hurter*, 2014):

- $f(\mathbf{r}_{ij}) \geq 0$ with $\mathbf{r}_{ij} \geq 0$
- $f(0) = 1$
- $\lim\limits_{\mathbf{r}_{ij} \to \infty} f(\mathbf{r}_{ij}) = 0$
- $-\infty < f'(\mathbf{r}_{ij}) < \infty$

- $f'(0) = 0$
- positive, monotonically decreasing

Typical functions that fulfill these requirements are:

$$\text{Model 1:} \qquad f(\mathbf{r}_{ij}) = \sigma_s^2 \cdot e^{-u\cdot|\mathbf{r}_i - \mathbf{r}_j|^n} \tag{5.8}$$

$$\text{Model 2:} \qquad f(\mathbf{r}_{ij}) = \sigma_s^2 \cdot \frac{1}{1 + u \cdot |\mathbf{r}_i - \mathbf{r}_j|^n} \tag{5.9}$$

$$\text{Model 3:} \qquad f(\mathbf{r}_{ij}) = \sigma_s^2 \cdot \frac{\sin(r_s \cdot |\mathbf{r}_i - \mathbf{r}_j|)}{r_s \cdot |\mathbf{r}_i - \mathbf{r}_j|} \tag{5.10}$$

$$\text{with} \quad u = 1/r_s^n \text{ and } n \in \mathbb{R}^+$$

The parameters $r_s$, $n$, and $\sigma_s^2$ are unknowns and consequently must either be set a priori or estimated empirically from the data. The choice of the parameter $n$ is not totally free, because the conditions listed above imply that the covariance matrix must be positive definite, which is the case as long as $n \in ]0, 2]$. More details regarding these constraints are given in *Rudin* (2011).

When the parameters of $f$ are determined empirically, the autocorrelation of the dataset is used for reference. As shown in *Hurter* (2014), estimating the correlation length is relatively robust with respect to sampling. An underestimated correlation length has the negative effect of yielding more noise in the collocation result. For large datasets a more

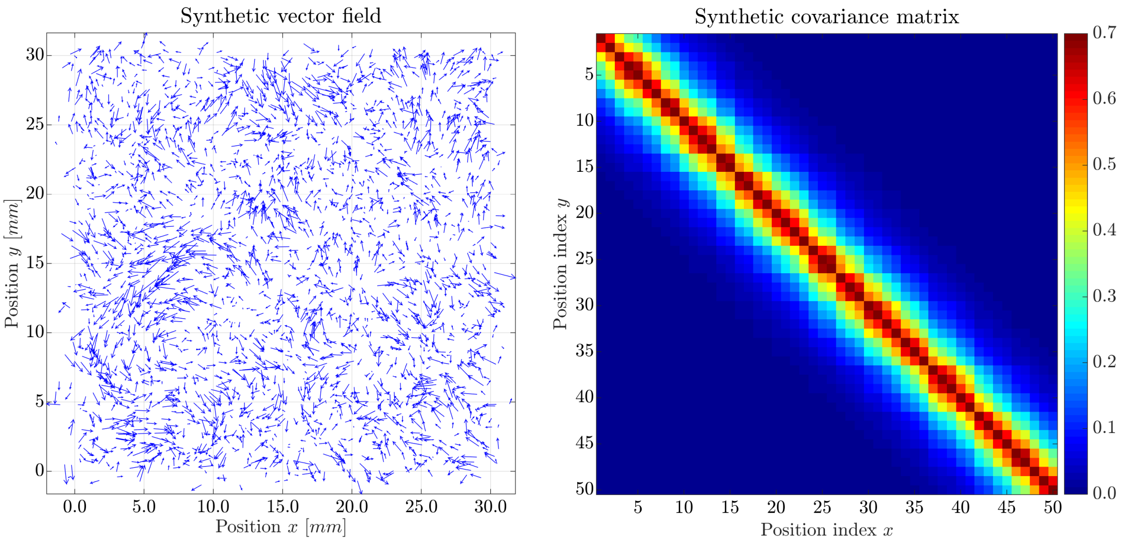


**Figure 5.1:** $(a)$ Synthetic vector field generated from a random variable $\sim \mathcal{N}(0; 1)$. $(b)$ shows the variance-covariance matrix with $f(\mathbf{r}) = 0.7 \cdot e^{-0.3\cdot|\mathbf{r}_i - \mathbf{r}_j|^{1.2}}$ being evaluated for all $|\mathbf{r}_i - \mathbf{r}_j|$. The position vector $\mathbf{r}$ was generated by weakly disturbing a regular grid (50 samples from 0 to 30 for both dimensions). More details about generating such a vector field can be found in the text.

robust estimate can be obtained, if binned correlation estimates $\hat{\mathbf{c}}(d_b)$ at different offsets are used. For each bin position, $d_b$ ($b$ being the distance index), the weighted average correlation over the distance range $k$ is computed as:

$$\mathbf{c}(d) = \mathbf{l}(\mathbf{r}_i) \cdot \mathbf{l}(\mathbf{r}_j) \qquad \text{for all} \quad |\mathbf{r}_i - \mathbf{r}_j| = d \tag{5.11}$$

$$\hat{\mathbf{c}}(d_b) = \frac{\sum\limits_{n=(b-1)\cdot k+1}^{b\cdot k \leq N} w_n \cdot \mathbf{c}(n)}{\sum\limits_{n=(b-1)\cdot k+1}^{b\cdot k \leq N} w_n} \tag{5.12}$$

with $N$ being the length of vector $\mathbf{c}(d)$ and $w$ the measurement weights. In other words, the estimated correlation at distance $d_b$ is proportional to the (weighted) average of all measurement products $\mathbf{l}(\mathbf{r}_i) \cdot \mathbf{l}(\mathbf{r}_j)$ with $d_{(b-1)\cdot k} < |\mathbf{r}_i - \mathbf{r}_j| \leq d_{b\cdot k}$. Depending on the complexity of the physical problem, the correlation vector $\hat{\mathbf{c}}$ can be multi-dimensional and $\hat{\mathbf{c}} \in \mathbb{R}^N$. For many practical problems, however, $\hat{\mathbf{c}}$ is defined in $\mathbb{R}^{+N}$ (as in Eqn.(5.12)).

**Table 5.1:** Summary of covariance parameter estimations using a synthetic velocity field with a given variance-covariance matrix, following $f(\mathbf{r}) = 0.7 \cdot e^{-0.4\cdot|\mathbf{r}_i - \mathbf{r}_j|^{1.3}}$, with $\sigma_n^2 = 0.4$. Each row shows the average result of 100 different realizations (see text for more details). The relative bin size states the number of position pairs within one bin relative to the total number of positions (here 900). Each estimate is given with its corresponding $1\sigma$ level.

| Relative bin size | $\sigma_s^2$ | $u$ | $n$ | absolute # bins |
|---|---|---|---|---|
| 8 | $0.850 \pm 0.032$ | $0.427 \pm 0.044$ | $1.533 \pm 0.074$ | 39 |
| 4 | $0.776 \pm 0.035$ | $0.469 \pm 0.049$ | $1.312 \pm 0.081$ | 76 |
| 2 | $0.727 \pm 0.031$ | $0.438 \pm 0.043$ | $1.343 \pm 0.071$ | 151 |
| 1 | $0.720 \pm 0.028$ | $0.423 \pm 0.039$ | $1.389 \pm 0.065$ | 300 |
| 1/2 | $0.703 \pm 0.028$ | $0.411 \pm 0.039$ | $1.383 \pm 0.065$ | 599 |
| 1/4 | $0.699 \pm 0.028$ | $0.419 \pm 0.039$ | $1.360 \pm 0.064$ | 1197 |
| 1/8 | $0.685 \pm 0.028$ | $0.405 \pm 0.039$ | $1.365 \pm 0.065$ | 2393 |
| 1/16 | $0.674 \pm 0.028$ | $0.413 \pm 0.039$ | $1.353 \pm 0.065$ | 4785 |
| 1/32 | $0.659 \pm 0.028$ | $0.435 \pm 0.039$ | $1.317 \pm 0.065$ | 9569 |
| 1/64 | $0.673 \pm 0.028$ | $0.487 \pm 0.040$ | $1.278 \pm 0.068$ | 19137 |

A synthetic dataset was generated to test the effect of data binning. Because collocation assumes normal data distribution, displacements were produced such that they follow $\sim \mathcal{N}(0; \mathbf{C})$. A realization of the vector $\mathbf{v}$ can be obtained as soon as the covariance matrix $\mathbf{C}$ is defined. Here, $\mathbf{C}$ was computed using model 1 (Eqn.(5.8)) with pre-defined parameters $\sigma_s^2$, $u$, and $n$, where $\mathbf{r}$ has the dimension of $900 \times 2$. Using the normalized eigenvectors $\mathbf{U}$ and corresponding eigenvalues $\mathbf{\Lambda}$ of $\mathbf{C}$, the vector $\mathbf{v}$ is computed as $\mathbf{v} = \mathbf{U} \cdot \sqrt{\mathbf{\Lambda}} \cdot \mathbf{y}$, with $\mathbf{y} \sim \mathcal{N}(0; 1)$ (*Guillaume*, 2014). Additional Gaussian noise ($\sim \mathcal{N}(0; \sigma_n^2)$) was added to $\mathbf{v}$. The result with $\sigma_s^2 = 0.70$, $u = 0.40$, $n = 1.30$, and $\sigma_n^2 = 0.40$ is summarized in Table 5.1. The numbers represent the average of 100 realizations and parameter estimations. An example of the 2-dimensional vector field with its (noise free) covariance matrix is shown in Fig. 5.1.

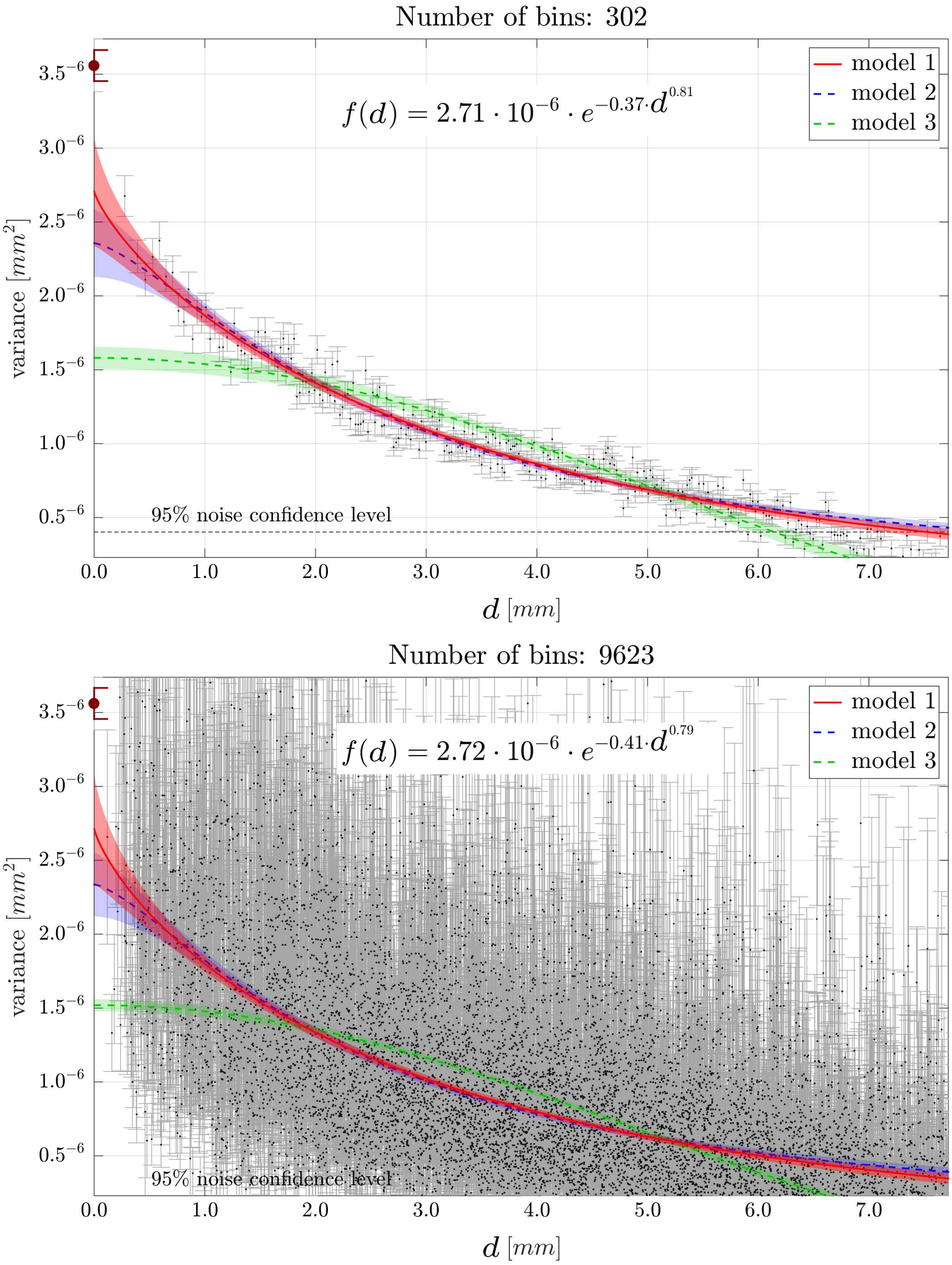


**Figure 5.2:** Empirical covariance function estimates. Black dots represent the mean covariance of binned data $\hat{\mathbf{c}}(d_b)$ with $1\sigma$ error bars in grey. The least-squares solution of the three analytical models (Eqns. (5.8) - (5.10)) are given in red, green, and blue, with the corresponding $3\sigma$ envelopes. $d$ represents the absolute point-to-point distance. The brown circle at $d = 0$ shows the mean variance with its $1\sigma$ standard deviation. The total variance is not used in the computation of the model parameters.

The covariogram (Eqn.(5.12)) of the synthetic vector field is computed using different bin sizes, whereas in every scenario, a least-squares adjustment is used to estimate the parameters of the covariance function (model 1, Eqn.(5.8)). Tabel 5.1 summarizes the result.

According to this test series, the optimal relative bin size is $\approx 1/2$ (i.e., the number of position pairs in one bin corresponds to about half the total number of measurement positions). In Fig. 5.2, the difference between well-binned (number of samples in each bin equals the number of measurements, $N$) and hardly binned (number of samples in each bin equals $1/32\,N$) correlation estimates are shown for a real dataset: displacement estimates of matched images features of a scene taken with different lenses at different image scales are used as measurements. A globally estimated transformation mapping the features from the patch image onto the template image was applied to compute the residuals. These residuals are then used to compute the covariogram (Eqn. (5.12)) that serves as input for estimating the empirical correlation functions (model 1 to 3, Eqns. 5.8 - 5.10). The resulting three curves in Fig. 5.2 were computed by least-squares adjustment, corresponding errors of $f$ were obtained using the error propagations laws. In Fig. 5.3, estimated parameters for every decreasing bin sizes are shown.

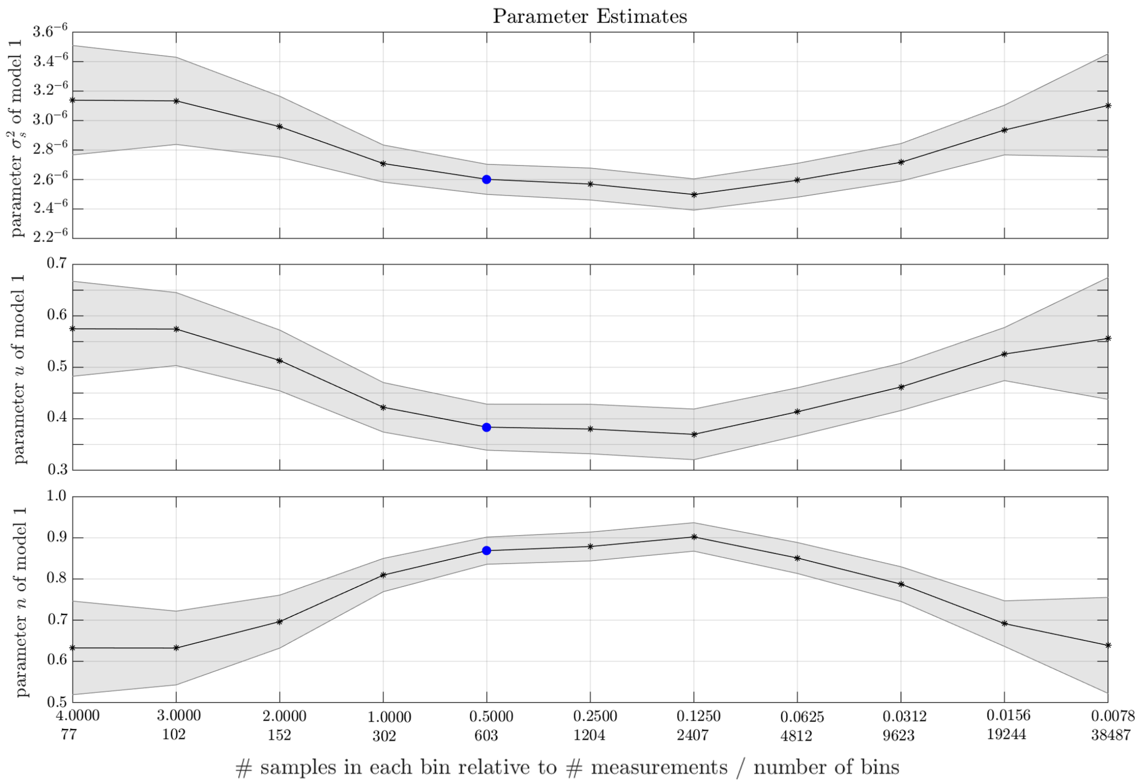


**Figure 5.3:** Parameters estimates of model 1 (Eqn.(5.8)) using different binning numbers. Error bar envelopes show the $1\sigma$ parameter standard deviation. The blue circle gives the optimal solution, i.e., minimum total parameter variance. Note, however, that estimates from relative bin numbers of 1 to 0.0312 show a very similar error distribution.

## 5.3 Adaptive Least-Squares Collocation

One of the limitations of least-squares collocation is the missing flexibility to handle sharp boundaries of different velocity fields. This is because the covariance function for a given set of parameters only depends on the coordinate differences of the measurement points (see Eqns. (5.8) - (5.10)). Also in case the covariance function has directional components (for example: $f(\mathbf{r}_{ij}) = \sigma_s^2 \cdot \frac{1}{1+u_1 \cdot \mathbf{r}_{xx}^2 + u_2 \cdot \mathbf{r}_{yy}^2 + u_3 \cdot \mathbf{r}_{xx} \cdot \mathbf{r}_{yy}}$ with $\mathbf{r}_{xx} = \mathbf{r}_{x_i} - \mathbf{r}_{x_j}$ being the distance between the positions $i$ and $j$ of the x-component covariance matrix $\mathbf{r}_{xx}$, it represents a relative measurement position relationship (i.e., independent of the absolute position). In this section, a modified least-squares collocation method is presented that iteratively decorrelates boundary areas. It is based on the empirical estimation of an anisotropic and inhomogeneous covariance function that is not known a priori.

Originally, this method was developed within the Swiss 4D project (*Egli et al.*, 2007) for GNSS data and levelling measurements. The principle is to decorrelate points with high strain rates by changing their measurement metric, i.e., the measurement coordinates are deformed such that the distance between points of high strain rate increases, thus yielding a smaller correlation. *Villiger* (2014) improved the original version of this method such that 3-dimensional GNSS and leveling measurements could by adaptively collocated.

In this section, the main steps of the theoretical background are given as well as a further improvement of the technique: the measurement epoch is also considered as a coordinate component making it possible to apply the metric dilation not only in space but also in time. The motivation to include this additional component can be found, when looking at time-depending phenomena like for example the flow behavior of (rock) glaciers, that is prone to rapid changes due to external triggering events (e.g., *Delaloye et al.*, 2010; *Wirz et al.*, 2013).

Sections 5.3.1 to 5.3.5 are based on the description found in *Egli et al.* (2007) and *Villiger* (2014), but formulated an arbitrary dimensionality. Notes on combining space and time with this procedure are given in Appendix D.

### 5.3.1 Principle

*Moritz* (1970) solves the least-squares collocation problem by dividing the solution vector $\mathbf{l}$ into a deterministic part $\mathbf{A} \cdot \mathbf{x}$ and a stochastic part $\hat{\mathbf{s}}$, where matrix $\mathbf{A} \cdot \mathbf{x}$ defines the a priori model. Because a priori models are always approximations, the deterministic trend model will suffer from errors that transfer into the stochastic part which, as a consequence, will introduce a bias and so lead to a signal $\mathbf{s}$ that is not fully stochastic. In most geodetic and geophysical processes, an exact a priori model is not available or known (e.g., *Villiger*, 2014). For many applications the covariance function of the residual signal is inhomogeneous (e.g., *Rummel and Schwarz*, 1977) and anisotropic (e.g., *Morrison*, 1977), the Adaptive Least-Squares Collocation (ALSC) approach introduces an inhomogeneous and anisotropic covariance function, defining $\mathbf{C}_{s_i s_i}$, that fulfills the same criteria as Eqn.(5.7):

$$\mathbf{C}_{s_i s_i} = \sigma_{s_i}^2 \cdot f(|\mathbf{r}_i - \mathbf{r}_j|, r_{s_i}) \tag{5.13}$$

with $r_{s_i}$ being the correlation length and $\sigma_{s_i}^2$ the variance of the inhomogeneous and anisotropic signal. The purpose is to correct the deterministic model for its bias.

Given that Eqn. (5.1) is the basic least-squares collocation approach, formulation (5.1) is extended by dividing the signal into a inhomogeneous part, $\hat{\mathbf{s}}_i$, and a homogeneous part, $\hat{\mathbf{s}}_h$. The total signal is then given by:

$$\mathbf{l}(\mathbf{r}) = \mathbf{A} \cdot \hat{\mathbf{x}} + \hat{\mathbf{s}}_i + \hat{\mathbf{s}}_h + \hat{\mathbf{n}} \tag{5.14}$$

with $\hat{\mathbf{s}}_i = \hat{\mathbf{s}}_i(\mathbf{r})$ and $\hat{\mathbf{s}}_h = \hat{\mathbf{s}}_h(\mathbf{r})$ defined as:

$$\hat{\mathbf{s}}_i = \mathbf{C}_{s_i s_i} \cdot (\mathbf{C}_{s_i s_i} + \mathbf{C}_{s_h s_h} + \mathbf{C}_{nn})^{-1} \cdot (\mathbf{l}(\mathbf{r}) - \mathbf{A} \cdot \hat{\mathbf{x}}) \tag{5.15a}$$

$$\hat{\mathbf{s}}_h = \mathbf{C}_{s_h s_h} \cdot (\mathbf{C}_{s_i s_i} + \mathbf{C}_{s_h s_h} + \mathbf{C}_{nn})^{-1} \cdot \mathbf{l}(\mathbf{r}) - \mathbf{A} \cdot \hat{\mathbf{x}}) \tag{5.15b}$$

The inhomogeneous and anisotropic covariance function (5.13) is not known a priori and must be determined iteratively. A good estimate is obtained if the inhomogeneity and anisotropy follow the local deformation rates. Thus the principle of the ALSC is to iteratively determine local deformation rates (i.e., strain rates) and based on these, estimate local metric deformations leading to $\mathbf{r}^{k-1} \to \mathbf{r}^k$, with $k$ being the iteration counter. This works under the assumption that strain rates are highest at motion boundaries (like for example rock glacier interfaces). High strain rates, therefore, indicate areas to be decorrelated, thus the inhomogeneous metric is iteratively dilated to achieve this effect. The collocation begins with the original metric, $\mathbf{r}^{k=0}$, to calculate $\hat{\mathbf{s}}_i$. Using the estimated deformation rates, a coordinate transformation is computed for the new metric, $\mathbf{r}^{k=0} \to \mathbf{r}^{k=1}$, so that the distance between points with high strain rates is is increased. Using this new metric, a better model for $\mathbf{C}_{\mathbf{s_i s_i}}$ can be computed and used again for a better strain rate estimate. The variance-covariance matrices after the $k$-th iteration hence are:

$$\mathbf{C}_{s_i s_i}^k = \sigma_{s_i}^2 \cdot f(|\mathbf{r}_i^{(k-1)} - \mathbf{r}_j^{(k-1)}|, r_{s_i}^k) \tag{5.16}$$

$$\mathbf{C}_{s_h s_h} = \sigma_{s_h}^2 \cdot f(|\mathbf{r}_i - \mathbf{r}_j|, r_{s_h}) \tag{5.17}$$

As indicated in Eqn. (5.17), the variance-covariance matrix of the stochastic signal $\hat{\mathbf{s}}_h$ is not effected by this so called dilation process (described below) and so it remains homogeneous.

### 5.3.2 Dilation Process

As described in the previous section, a process called dilation is needed to achieve an 'uncoupling' of regions where high deformation rates indicate a weak correlation. This increase in point-to-point distance follows a derivable vector function (*Egli et al.*, 2007; *Villiger*, 2014) that defines the coordinate transformation for the new metric:

$$\mathbf{r}^{tr} = \mathbf{r} + \mathbf{D}(\mathbf{r}, \mathbf{r}_0, \mathbf{u}_1, .., \mathbf{u}_s, e_1, .., e_s, \gamma) \tag{5.18}$$

with $\mathbf{r}^{tr}$ being the transformed coordinate, $\mathbf{r}$ the original coordinate, $\mathbf{r}_0$ the center of dilation, $\mathbf{u}_1, .., \mathbf{u}_s$ the orientation of the orthogonal unit vectors in all $s$ dimensions, the scaling coefficients $e_1, .., e_s$, and the dilation length $\gamma$.

As explained in *Egli et al.* (2007), this function has to fulfill two properties:

$$\lim_{|\mathbf{r}-\mathbf{r}_0|/\gamma \to 0} \mathbf{r}^{tr} - \mathbf{r}_0 = \sum_{i=1}^{s} e_i \cdot < \mathbf{r} - \mathbf{r}_0, \mathbf{u}_i > \tag{5.19}$$

$$\lim_{|\mathbf{r}-\mathbf{r}_0|/\gamma \to \infty} \mathbf{r}^{tr} - \mathbf{r}_0 = 0 \tag{5.20}$$

with $< \mathbf{a}, \mathbf{b} >$ representing the scalar product between the vectors $\mathbf{a}$ and $\mathbf{b}$. A simple formulation that meets these requirements is:

$$\mathbf{D}(\mathbf{r}, \mathbf{r}_0, \mathbf{u}_1, .., \mathbf{u}_s, e_1, .., e_s, \gamma) = \triangle(|\mathbf{r} - \mathbf{r}_0|, \gamma) \sum_{l=1}^{s} e_l \cdot < \mathbf{r} - \mathbf{r}_0, \mathbf{u}_l > \cdot \mathbf{u}_l \tag{5.21}$$

$$\text{with} \quad \triangle(r, \gamma) = \frac{\gamma}{r} \tanh\left(\frac{r}{\gamma}\right) \tag{5.22}$$

Function $\mathbf{D}$ thus describes the dilation caused by a single source (located at $\mathbf{r}_0$). The total coordinate transformation of $N$ dilation sources ($\mathbf{r}_0 = \mathbf{r}_1, .., \mathbf{r}_N$) is obtained by their superposition:

$$\mathbf{r}^{tr} = \mathbf{r} + \sum_{k=1}^{N} \mathbf{D}(\mathbf{r}, \mathbf{r}_k, \mathbf{u}_{k1}, .., \mathbf{u}_{ks}, e_{k1}, .., e_{ks}, \gamma) \tag{5.23}$$

To apply Eqn. (5.23), dilation directions $\mathbf{u}_{k1}, .., \mathbf{u}_{ks}$ and dilation stretching factors $e_{k1}, .., e_{ks}$ have to be determined.

### 5.3.3 Estimating Dilation Parameters

The gradient of the coordinate transformation described in Eqn. (5.23) is given by the directional derivative of $\mathbf{D}$. In direction of $\mathbf{n}$ this is given as:

$$\frac{\partial}{\partial \mathbf{n}} \mathbf{D}(\mathbf{r}, \mathbf{r}_0, \mathbf{u}_1, .., \mathbf{u}_s, e_1, .., e_s, \gamma) = \sum_{l=1}^{s} e_l < \mathbf{G}(\mathbf{r} - \mathbf{r}_0, \mathbf{n}, \gamma), \mathbf{u}_l > \mathbf{u}_l \tag{5.24}$$

with

$$\mathbf{G}(\mathbf{r}, \mathbf{n}, \gamma) =< \mathbf{r}, \mathbf{n} > \frac{\mathbf{r}}{r^2} \operatorname{sech}^2\left(\frac{r}{\gamma}\right) + \frac{\gamma}{r}\Big[\, \mathbf{n} - < \mathbf{r}, \mathbf{n} > \frac{\mathbf{r}}{r^2}\Big] \tanh\left(\frac{r}{\gamma}\right) \tag{5.25}$$

$$r = |\mathbf{r}_i - \mathbf{r}_j|$$

By considering Eqns. (5.24) and (5.25), and following *Egli et al.* (2007) in combination with *Villiger* (2014), the derivative of Eqn. (5.23) for a point $\mathbf{r}_j$ with $1 \leq j \leq N$ along the orthogonal directions $\mathbf{d}_{j1}, .., \mathbf{d}_{js}$ is given by:

$$\sum_{k=1}^{N}\sum_{l=1}^{s} e_{kl} < \mathbf{G}_{kj1}, \mathbf{u}_{kl} > \mathbf{u}_{kl} = d_{j1}\mathbf{d}_{j1} \tag{5.26a}$$

$$\sum_{k=1}^{N}\sum_{l=1}^{s} e_{kl} < \mathbf{G}_{kj2}, \mathbf{u}_{kl} > \mathbf{u}_{kl} = d_{j2}\mathbf{d}_{j2} \tag{5.26b}$$

$$\vdots$$

$$\sum_{k=1}^{N}\sum_{l=1}^{s} e_{kl} < \mathbf{G}_{kjs}, \mathbf{u}_{kl} > \mathbf{u}_{kl} = d_{js}\mathbf{d}_{js} \tag{5.26c}$$

where $\mathbf{G}_{kjs} = \mathbf{G}(\mathbf{r}_j - \mathbf{r}_k, \mathbf{d}_{js}, \gamma)$. When replacing the unit vectors $\mathbf{u}_{kl}$ with $\mathbf{d}_{kl}$ and building the scalar product of all components (Eqns. (5.26)), the result is:

$$\sum_{k=1}^{N}\sum_{l=1}^{s} e_{kl} < \mathbf{G}_{kjs}, \mathbf{d}_{kl} >< \mathbf{d}_{kl}, \mathbf{d}_{jm} > = d_{js}\delta_{ml} \tag{5.27}$$

with $m = 1, 2, .., s$ and the Kronecker delta $\delta_{ml}$. For any two points $\mathbf{r}_j$, $\mathbf{r}_i$, the following assumptions can be applied:

(1) $< \mathbf{d}_{kl}, \mathbf{d}_{jm} > \approx 0$ for $l \neq m$ and $|\mathbf{r}_j - \mathbf{r}_k| < \gamma$

(2) $\mathbf{G}_{kjl}\mathbf{d}_{km} \approx 0$ for $l \neq m$ and $|\mathbf{r}_j - \mathbf{r}_k| < \gamma$

(3) $\mathbf{G}_{kjl} \approx 0$ for $|\mathbf{r}_j - \mathbf{r}_k| \gg \gamma$

Considering these simplifications, the system of equations can be written as:

$$\sum_{k=1}^{N}\sum_{l=1}^{s} e_{kl} < \mathbf{G}_{kj1}, \mathbf{d}_{kl} >< \mathbf{d}_{kl}, \mathbf{d}_{j1} > = d_{j1} \tag{5.28a}$$

$$\sum_{k=1}^{N}\sum_{l=1}^{s} e_{kl} < \mathbf{G}_{kj2}, \mathbf{d}_{kl} >< \mathbf{d}_{kl}, \mathbf{d}_{j2} > = d_{j2} \tag{5.28b}$$

$$\vdots$$

$$\sum_{k=1}^{N}\sum_{l=1}^{s} e_{kl} < \mathbf{G}_{kjs}, \mathbf{d}_{kl} >< \mathbf{d}_{kl}, \mathbf{d}_{js} > = d_{js} \tag{5.28c}$$

that is a set of $s \cdot N$ linear equations with $s \cdot N$ unknowns $e_{kl}$. $\mathbf{d}_{kl}$ and $d_{kl}$ are the eigenvectors and eigenvalues of the strain tensor for point $k$. The system of equations can be rearranged such that the unknown scaling values $e_{kl}$ are obtained as:

$$\begin{pmatrix} \mathbf{e}_1 \\ \mathbf{e}_2 \\ \vdots \\ \mathbf{e}_s \end{pmatrix} = \begin{pmatrix} \mathbf{M}_{11} & \mathbf{M}_{12} & \cdots & \mathbf{M}_{1s} \\ \mathbf{M}_{21} & \mathbf{M}_{22} & \cdots & \mathbf{M}_{2s} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{M}_{s1} & \mathbf{M}_{s2} & \cdots & \mathbf{M}_{ss} \end{pmatrix}^{-1} \cdot \begin{pmatrix} \mathbf{v}_1 \\ \mathbf{v}_2 \\ \vdots \\ \mathbf{v}_s \end{pmatrix} \tag{5.29}$$

with $[\mathbf{M}_{lm}]_{kj} = (\mathbf{G}_{kjl}\mathbf{d}_{km})(\mathbf{d}_{km}\mathbf{d}_{jl})$, $[\mathbf{e}_l]_k = e_{kl}$ and $[\mathbf{v}_l]_j = d_{jl}$. Thus, once the strain tensors at the dilation point sources are known, the deformation scaling rates can be computed.

The strain rate tensor is a product of the velocity gradient field $\nabla\hat{\mathbf{s}}_i'$, that, in the least-squares collocation for the interpolated inhomogeneous signal, obtained by:

$$\nabla\hat{\mathbf{s}}_i' = \nabla\mathbf{C}_{s_i's_i} \cdot (\mathbf{C}_{s_is_i} + \mathbf{C}_{s_hs_h} + \mathbf{C}_{nn})^{-1} \cdot (\mathbf{l}(\mathbf{r}) - \mathbf{A}\cdot\hat{\mathbf{x}}) = \nabla\mathbf{C}_{s_i's_i} \cdot \mathbf{k} \tag{5.30}$$

$$\mathbf{k} = (\mathbf{C}_{s_is_i} + \mathbf{C}_{s_hs_h} + \mathbf{C}_{nn})^{-1} \cdot (\mathbf{l}(\mathbf{r}) - \mathbf{A}\cdot\hat{\mathbf{x}}) \tag{5.31}$$

In case the design matrix $\mathbf{A}$ describes a deterministic part other than a offset, $\nabla\mathbf{A}$ also has to be considered (i.e., added) for computing the strain rates of the total trend (i.e., deterministic and inhomogeneous part, $\tau = \mathbf{A}\cdot\hat{\mathbf{x}} + \hat{\mathbf{s}}_i$, compare Section 5.1.2):

$$\nabla\tau = \nabla\mathbf{A}\cdot\hat{\mathbf{x}} + \nabla\hat{\mathbf{s}}_i \tag{5.32}$$

In the following Section (5.3.4), $\nabla\mathbf{A}$ is assumed to be zero. However adding this component is trivial because its contribution only depends on the estimated parameters of the deterministic model and the initial interpolation positions.

### 5.3.4 Strain Rate Tensor

The general definition of the strain rate tensor $\epsilon_k$ for a point $\mathbf{r}_k$ with velocity vector $\mathbf{s}_{i,k}$ in three dimensions is:

$$\epsilon_k = \frac{1}{2}\left((\nabla\mathbf{s}_{i,k})^T + \nabla\mathbf{s}_{i,k}\right) \tag{5.33}$$

$$\text{with} \quad \nabla\mathbf{s}_i = \begin{pmatrix} \nabla s_{i_x} \\ \nabla s_{i_y} \\ \nabla s_{i_z} \end{pmatrix} = \begin{pmatrix} \frac{\partial s_{i_x}}{\partial x} & \frac{\partial s_{i_x}}{\partial y} & \frac{\partial s_{i_x}}{\partial z} \\ \frac{\partial s_{i_y}}{\partial x} & \frac{\partial s_{i_y}}{\partial y} & \frac{\partial s_{i_y}}{\partial z} \\ \frac{\partial s_{i_z}}{\partial x} & \frac{\partial s_{i_z}}{\partial y} & \frac{\partial s_{i_z}}{\partial z} \end{pmatrix} \tag{5.34}$$

During the iterative process of the ALSC, the trend metric changes. This does not have a direct impact on the velocity estimations (Eqns. (5.14) and (5.15)) but on the strain rate tensor. This is because the dilation process deforms the coordinate grid (originally orthogonal unit vector directions) such that the distances and directions are being deformed. Thus for estimating the strain rates in the undisturbed grid, the directional derivatives, given by the following expression, have to be considered:

$$\frac{\partial\hat{\mathbf{s}}_i}{\partial x} = \frac{\partial x^{tr}}{\partial x}\frac{\partial\hat{\mathbf{s}}_i}{\partial x^{tr}} + \frac{\partial y^{tr}}{\partial x}\frac{\partial\hat{\mathbf{s}}_i}{\partial y^{tr}} + \frac{\partial z^{tr}}{\partial x}\frac{\partial\hat{\mathbf{s}}_i}{\partial z^{tr}} \tag{5.35a}$$

$$\frac{\partial\hat{\mathbf{s}}_i}{\partial y} = \frac{\partial x^{tr}}{\partial y}\frac{\partial\hat{\mathbf{s}}_i}{\partial x^{tr}} + \frac{\partial y^{tr}}{\partial y}\frac{\partial\hat{\mathbf{s}}_i}{\partial y^{tr}} + \frac{\partial z^{tr}}{\partial y}\frac{\partial\hat{\mathbf{s}}_i}{\partial z^{tr}} \tag{5.35b}$$

$$\frac{\partial\hat{\mathbf{s}}_i}{\partial z} = \frac{\partial x^{tr}}{\partial z}\frac{\partial\hat{\mathbf{s}}_i}{\partial x^{tr}} + \frac{\partial y^{tr}}{\partial z}\frac{\partial\hat{\mathbf{s}}_i}{\partial y^{tr}} + \frac{\partial z^{tr}}{\partial z}\frac{\partial\hat{\mathbf{s}}_i}{\partial z^{tr}} \tag{5.35c}$$

with (compare Eqn. (5.30)):

$$\frac{\partial \hat{\mathbf{s}}_i}{\partial x^{tr}} = \frac{\partial \mathbf{C}_{s_i' s_i}}{\partial x^{tr}} \cdot \mathbf{k} \tag{5.36a}$$

$$\frac{\partial \hat{\mathbf{s}}_i}{\partial y^{tr}} = \frac{\partial \mathbf{C}_{s_i' s_i}}{\partial y^{tr}} \cdot \mathbf{k} \tag{5.36b}$$

$$\frac{\partial \hat{\mathbf{s}}_i}{\partial z^{tr}} = \frac{\partial \mathbf{C}_{s_i' s_i}}{\partial z^{tr}} \cdot \mathbf{k} \tag{5.36c}$$

Practically, the correction terms $\frac{\partial x^{tr}}{\partial x}$, $\frac{\partial y^{tr}}{\partial x}$, $\frac{\partial x^{tr}}{\partial y}$, etc., for the $(k+1)$-th iteration are determined by multiplying the current estimates with the previous ones. For the first component $\frac{\partial x^{tr}}{\partial x}$, this reads as:

$$\frac{\partial {x^{tr}}^{k+1}}{\partial x} = \frac{\partial {x^{tr}}^{k-1}}{\partial x} \cdot \frac{\partial {x^{tr}}^{k}}{\partial x} \tag{5.37}$$

The eigenvectors $\mathbf{d}_{jl}$ and the corresponding eigenvalues $\lambda_{jl}$ (used to estimate the dilation parameters, Eqn. (5.29)) are obtained using an eigenvalue decomposition on the corrected strain rate tensor.

### 5.3.5 Linking Gradient Field and Transformation Parameters

A suitable coordinate transformation is obtained if the dilation directions, given by the eigenvectors $\mathbf{d}_{jl}$, coincide with the directional derivative $|\mathbf{u}\nabla\hat{\mathbf{s}}_i|$. In addition, the amount of dilation, given by the eigenvalues $d_{jl}$ should be proportional to $|\mathbf{d}_{jl}\nabla\hat{\mathbf{s}}_i|$. These conditions assure, that the main dilation component is acting in the direction of the maximum gradient and is proportional to the gradient itself. To prevent the situation for arbitrary dilations, a scaling parameter $\lambda_{max}$ is introduced, defining the maximum allowed dilation between any two points. Applying an eigenvalue decomposition of the strain rate tensor, yielding the eigenvectors $\mathbf{d}_{jl}$ and corresponding eigenvalues $\lambda_{jl}$, the dilation scales are obtained by:

$$d_{jl} = \lambda_{max} \cdot g_{max}^{-1} \cdot |\lambda_{jl}| \tag{5.38}$$

with $g_{max}$ being a truncation value so that:

$$d_{jl} = \begin{cases} 1 & \text{if } g_{max}^{-1} \cdot \lambda_{jl} > 1 \\ \lambda_{max} \cdot g_{max}^{-1} \cdot |\lambda_{jl}| & \text{otherwise} \end{cases} \tag{5.39}$$

The truncation parameter $g_{max}$ is used to define an upper threshold for the estimated eigenvalues. In cases of multiple trend boundaries with highly different velocity components (for example the velocity gradient at the first boundary is by a factor of 10 larger than at the second boundary) the estimated dilation scales would also show a large difference and thus the dominated dilation would only occur at the highest strain rate areas. By introducing a clipping parameter, the eigenvalues are clipped (as shown in Eqn. (5.39)) so that blocks eventually are dilated by an equal amount (compare Fig. 5.4).

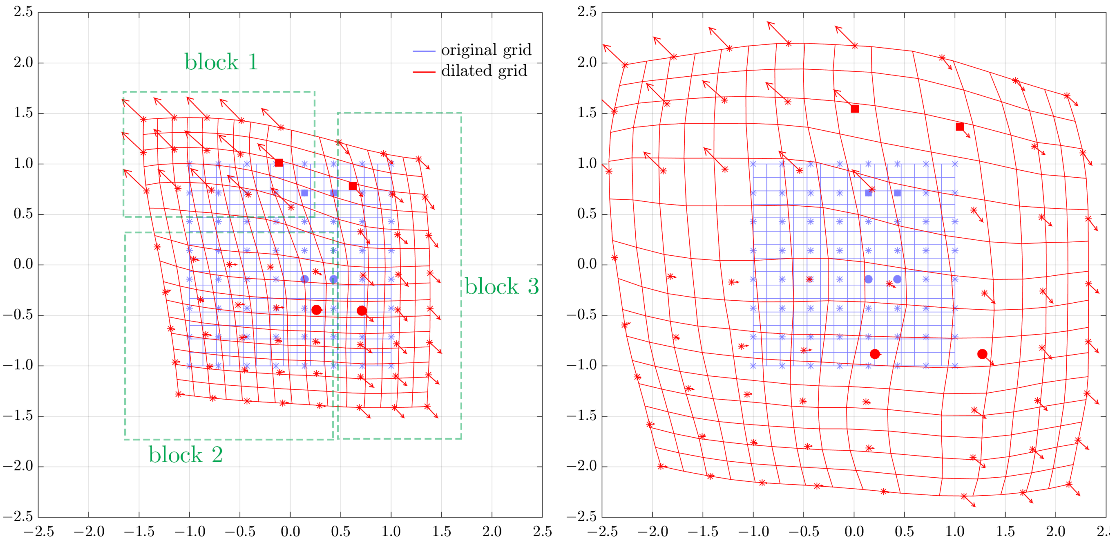


**Figure 5.4:** Comparison of different $g_{max}$ thresholds used for the dilation scale estimation after three iterations. Left plot: $g_{max}$ was set to the maximum gradient in the first iteration ($g_{max}$ = 16.2). Right plot: equal settings as in the left plot but $g_{max} = 3$. Arrows indicate relative velocities, $\mathbf{l}(\mathbf{r})$, of the three blocks. Due to the given velocities, the gradient between block 1 and block 3 is larger than between block 2 and block 3. Thus in the left dilation process, dilation is more pronounced between block 1 and 3: the ratio of the distance between the positions marked as squares and the distance between the positions marked as circles is 1.72. The same ratio in the right plot is 0.99

### 5.3.6 Estimating Optimal ALSC Components

The general problem of the ALSC technique is to find the optimal parameters $r^k_{s_i}$, $r_{s_h}$, $\sigma^2_{s_i}$, $\sigma^2_{s_h}$, $\lambda_{max}$, and eventually also $g_{max}$ without a detailed a priori understanding of the velocity field. The assumption of an uncorrelated residual $\hat{\mathbf{n}}$ (Eqn. (5.14)) can be obtained having different signal parts $\hat{\mathbf{s}}_i$ and $\hat{\mathbf{s}}_h$. Thus the following strategy is proposed: first it is assumed that $\sigma^2_{s_h}$ is zero during the ALSC iterations, yielding a residual part $\hat{\mathbf{n}}_1 = \hat{\mathbf{s}}_h + \hat{\mathbf{n}}_2$ that eventually has some stochastically correlated residual component. In the second step the parameters for $\mathbf{C}_{s_h s_h}$ are determined either using an empirical estimate of the correlation function (see Section 5.2) or by iterating such that the norm of the autocorrelation of the residual $\hat{\mathbf{n}}_2$ gets minimal. This, however, solves only for one part of the unknown parameters. The initial trend correlation length $r^{k=0}_{s_i}$, the dilation length $\gamma$, the truncation value $g_{max}$, and the ratio of the variance components of the inhomogeneous and homogeneous signal part, $\sigma^2_{s_i}$ and $\sigma^2_{s_h}$, are - at this stage - determined by the trial and error principles. The initial trend correlation length is typically chosen such that it corresponds to the expected trend deformation (e.g. *Villiger*, 2014).

## 5.4 Variance-Covariance Matrices of Collocated Quantities

So far, the interest was to derive collocated estimates of noisy measurements or derived quantities thereof. This section gives the respective formulations the corresponding variance-covariance matrices used to estimate the uncertainty of the collocated quantities.

### 5.4.1 Variance-Covariance of LSC Components

A derivation for the variance-covariance matrix of the collocated signal is foung, e.g., in *Wirth* (1990). As a detailed derivation of the uncertainties for the total signal of the ALSC procedure is given in Section 5.4.2, only the main results are listed here.

Following the definition given in Eqn. 5.4, the variance-covariance matrices, $\mathbf{Q}$, of the respective components are:

$$\mathbf{Q}_{\hat{l}'\hat{l}'} = \mathbf{C}_{s's'} - \mathbf{C}_{s's} \cdot \mathbf{D}^{-1} \cdot \mathbf{C}^T_{s's} + \mathbf{M} \cdot \mathbf{Q}_{\hat{x}\hat{x}} \cdot \mathbf{M}^T \tag{5.40}$$

$$\text{with} \quad \mathbf{M} = \mathbf{C}_{s's} \cdot \mathbf{D}^{-1} \cdot \mathbf{A} - \mathbf{A}'$$

$$\mathbf{Q}_{\hat{x}\hat{x}} = \left(\mathbf{A}^T\mathbf{D}^{-1}\mathbf{A}\right)^{-1} \tag{5.41}$$

### 5.4.2 Variance-Covariance of ALSC Components

To estimate the uncertainties of the total signal obtained by the adaptive collocation approach, a formulation of $\sigma^2_{\hat{\mathbf{l}}'}$ has to be found by the principles of error propagation. A discussion on error propagation for LSC is given for example in *Wirth* (1990) or *Moritz* (1973), and the main results are presented in Section 5.4.1. Following the notation given in *Wirth* (1990), the estimated error of the interpolated signal $\hat{\mathbf{l}}' = \mathbf{A}'\hat{\mathbf{x}} + \hat{\mathbf{s}}'_i + \hat{\mathbf{s}}'_h$, with error matrix $\mathbf{Q}_{\hat{l}'\hat{l}'} = E\langle\epsilon_{\hat{l}'}\epsilon^T_{\hat{l}'}\rangle$, and $\epsilon_{\hat{l}'} = \hat{\mathbf{l}}' - \overline{\mathbf{l}}'$, $\overline{\mathbf{l}}'$ being the error free solution, is:

$$\begin{aligned}
\epsilon_{\hat{l}'} &= \hat{\mathbf{l}}' - \overline{\mathbf{l}}' \\
&= \mathbf{A}'\hat{\mathbf{x}} + \hat{\mathbf{s}}'_i + \hat{\mathbf{s}}'_h - \left(\mathbf{A}'\overline{\mathbf{x}} + \overline{\mathbf{s}}'_i + \overline{\mathbf{s}}'_h\right) \\
&= \left(\mathbf{A}'\mathbf{G} + \mathbf{L}_i + \mathbf{L}_h\right)\left(\mathbf{A}\overline{\mathbf{x}} + \overline{\mathbf{z}}\right) - \left(\mathbf{A}'\overline{\mathbf{x}} + \overline{\mathbf{s}}'_i + \overline{\mathbf{s}}'_h\right) \\
&= \left(\mathbf{A}'\mathbf{G}\mathbf{A} + \mathbf{L}_i\mathbf{A} + \mathbf{L}_h\mathbf{A}\right)\overline{\mathbf{x}} + \left(\mathbf{A}'\mathbf{G} + \mathbf{L}_i + \mathbf{L}_h\right)\overline{\mathbf{z}} - \left(\mathbf{A}'\overline{\mathbf{x}} + \overline{\mathbf{s}}'_i + \overline{\mathbf{s}}'_h\right)
\end{aligned} \tag{5.42}$$

with

$$\overline{\mathbf{z}} = \overline{\mathbf{s}}_i + \overline{\mathbf{s}}_h + \overline{\mathbf{n}} \tag{5.43}$$

$$\hat{\mathbf{x}} = \left(\mathbf{A}^T\mathbf{D}^{-1}\mathbf{A}\right)^{-1}\mathbf{A}^T\mathbf{D}^{-1}\mathbf{l} = \mathbf{Q}_{\hat{x}\hat{x}}\mathbf{A}^T\mathbf{D}^{-1}\mathbf{l} = \mathbf{G}\mathbf{l} \tag{5.44}$$

$$\begin{aligned}
\hat{\mathbf{s}}'_i &= \mathbf{C}_{s'_i s_i}\mathbf{D}^{-1} \cdot (\mathbf{l} - \mathbf{A} \cdot \hat{\mathbf{x}}) \\
&= \mathbf{H}_i \cdot (\mathbf{l} - \mathbf{A}\mathbf{G}\mathbf{l}) = \mathbf{H}_i \cdot (\mathbf{I} - \mathbf{A}\mathbf{G})\mathbf{l} = \mathbf{L}_i \cdot \mathbf{l}
\end{aligned} \tag{5.45}$$

$$\begin{aligned}
\hat{\mathbf{s}}'_h &= \mathbf{C}_{s'_h s_h}\mathbf{D}^{-1} \cdot (\mathbf{l} - \mathbf{A} \cdot \hat{\mathbf{x}}) \\
&= \mathbf{H}_h \cdot (\mathbf{l} - \mathbf{A}\mathbf{G}\mathbf{l}) = \mathbf{H}_h \cdot (\mathbf{I} - \mathbf{A}\mathbf{G})\mathbf{l} = \mathbf{L}_h \cdot \mathbf{l}
\end{aligned} \tag{5.46}$$

noting that

$$\mathbf{GA} = \mathbf{I}, \quad \mathbf{I} = \text{identity matrix}$$
$$\mathbf{L}_i\mathbf{A} = \mathbf{L}_h\mathbf{A} = 0$$

leads to

$$\epsilon_{\hat{l}'} = (\mathbf{A}'\mathbf{G} + \mathbf{L}_i + \mathbf{L}_h)\bar{\mathbf{z}} - \bar{\mathbf{s}}_i' - \bar{\mathbf{s}}_h' = \mathbf{K}\bar{\mathbf{z}} - \bar{\mathbf{s}}_i' - \bar{\mathbf{s}}_h' \tag{5.47}$$

$$\begin{aligned}\epsilon_{\hat{l}'}\epsilon_{\hat{l}'}^T &= (\mathbf{K}\bar{\mathbf{z}} - \bar{\mathbf{s}}_i' - \bar{\mathbf{s}}_h') \cdot (\bar{\mathbf{z}}^T\mathbf{K}^T - \bar{\mathbf{s}}_i'^T - \bar{\mathbf{s}}_h'^T) \\ &= \mathbf{K}\bar{\mathbf{z}}\bar{\mathbf{z}}^T\mathbf{K}^T - \mathbf{K}\bar{\mathbf{z}}\bar{\mathbf{s}}_i'^T - \mathbf{K}\bar{\mathbf{z}}\bar{\mathbf{s}}_h'^T - \bar{\mathbf{s}}_i'\bar{\mathbf{z}}^T\mathbf{K}^T \\ &\quad + \bar{\mathbf{s}}_i'\bar{\mathbf{s}}_i'^T + \bar{\mathbf{s}}_i'\bar{\mathbf{s}}_h'^T - \bar{\mathbf{s}}_h'\bar{\mathbf{z}}^T\mathbf{K}^T + \bar{\mathbf{s}}_h'\bar{\mathbf{s}}_i'^T + \bar{\mathbf{s}}_h'\bar{\mathbf{s}}_h'^T\end{aligned} \tag{5.48}$$

Because there is (by definition) no correlation between the $\bar{\mathbf{s}}_i'$ and $\bar{\mathbf{s}}_h'$, the corresponding error matrices are given as $E\langle\bar{\mathbf{s}}_i'\bar{\mathbf{s}}_h'^T\rangle = E\langle\bar{\mathbf{s}}_h'\bar{\mathbf{s}}_i'^T\rangle = 0$. All other error matrices are defined as:

$$E\langle\bar{\mathbf{s}}_i'\bar{\mathbf{s}}_i'^T\rangle = \mathbf{C}_{s_i's_i'} \tag{5.49}$$

$$E\langle\bar{\mathbf{s}}_h'\bar{\mathbf{s}}_h'^T\rangle = \mathbf{C}_{s_h's_h'} \tag{5.50}$$

$$E\langle\bar{\mathbf{z}}\bar{\mathbf{z}}^T\rangle = \mathbf{C}_{zz} \tag{5.51}$$

$$\begin{aligned}E\langle\bar{\mathbf{s}}_i'\bar{\mathbf{z}}^T\rangle &= E\langle\bar{\mathbf{s}}_i'(\bar{\mathbf{s}}_i + \bar{\mathbf{s}}_h + \bar{\mathbf{n}})^T\rangle = E\langle\bar{\mathbf{s}}_i'\bar{\mathbf{s}}_i^T\rangle + E\langle\bar{\mathbf{s}}_i'\bar{\mathbf{s}}_h^T\rangle + E\langle\bar{\mathbf{s}}_i'\bar{\mathbf{n}}^T\rangle \\ &= E\langle\bar{\mathbf{s}}_i'\bar{\mathbf{s}}_i^T\rangle = \mathbf{C}_{s_i's_i}\end{aligned} \tag{5.52}$$

$$E\langle\bar{\mathbf{z}}\bar{\mathbf{s}}_i'^T\rangle = \mathbf{C}_{s_is_i'} = \mathbf{C}_{s_i's_i}^T \tag{5.53}$$

so that:

$$\begin{aligned}E\langle\epsilon_{\hat{l}'}\epsilon_{\hat{l}'}^T\rangle = \mathbf{Q}_{\hat{l}'\hat{l}'} &= \mathbf{C}_{s_i's_i'} + \mathbf{C}_{s_h's_h'} - \mathbf{K}\cdot(\mathbf{C}_{s_i's_i}^T + \mathbf{C}_{s_h's_h}^T) \\ &\quad - (\mathbf{C}_{s_i's_i} + \mathbf{C}_{s_h's_h})\cdot\mathbf{K}^T + \mathbf{K}\mathbf{C}_{zz}\mathbf{K}^T\end{aligned} \tag{5.54}$$

When now looking at the individual terms of Eqn. (5.54), noting that $(\mathbf{A}^T\mathbf{D}^{-1}\mathbf{A})^{-1} = \mathbf{Q}_{\hat{x}\hat{x}} = \mathbf{Q}_{\hat{x}\hat{x}}^T$, their expansion yields:

$$\begin{aligned}\mathbf{K}\cdot(\mathbf{C}_{s_i's_i}^T + \mathbf{C}_{s_h's_h}^T) &= (\mathbf{A}'\mathbf{G} + \mathbf{L}_i + \mathbf{L}_h)\cdot(\mathbf{C}_{s_i's_i}^T + \mathbf{C}_{s_h's_h}^T) \\ &= (\mathbf{A}'\mathbf{Q}_{\hat{x}\hat{x}}\mathbf{A}^T\mathbf{D}^{-1} + \mathbf{L}_i + \mathbf{L}_h)\cdot(\mathbf{C}_{s_i's_i}^T + \mathbf{C}_{s_h's_h}^T) \\ &= (\mathbf{A}'\mathbf{Q}_{\hat{x}\hat{x}}\mathbf{A}^T)\cdot(\mathbf{H}_i^T + \mathbf{H}_h^T) + (\mathbf{L}_i + \mathbf{L}_h)\cdot(\mathbf{C}_{s_i's_i}^T + \mathbf{C}_{s_h's_h}^T)\end{aligned} \tag{5.55}$$

$$\begin{aligned}(\mathbf{C}_{s_i's_i} + \mathbf{C}_{s_h's_h})\cdot\mathbf{K}^T &= (\mathbf{C}_{s_i's_i} + \mathbf{C}_{s_h's_h})\cdot(\mathbf{D}^{-1}\mathbf{A}\mathbf{Q}_{\hat{x}\hat{x}}\mathbf{A}'^T + \mathbf{L}_i^T + \mathbf{L}_h^T) \\ &= (\mathbf{H}_i + \mathbf{H}_h)\cdot(\mathbf{A}\mathbf{Q}_{\hat{x}\hat{x}}\mathbf{A}'^T) + (\mathbf{C}_{s_i's_i} + \mathbf{C}_{s_h's_h})\cdot(\mathbf{L}_i^T + \mathbf{L}_h^T)\end{aligned} \tag{5.56}$$

$$\begin{aligned}\mathbf{K}\mathbf{C}_{zz}\mathbf{K}^T &= (\mathbf{A}'\mathbf{G} + \mathbf{L}_i + \mathbf{L}_h)\cdot\mathbf{C}_{zz}\cdot(\mathbf{G}^T\mathbf{A}'^T + \mathbf{L}_i^T + \mathbf{L}_h^T) \\ &= \mathbf{A}'\mathbf{G}\mathbf{D}\mathbf{G}^T\mathbf{A}'^T + \mathbf{A}'\mathbf{G}\mathbf{D}\mathbf{L}_i^T + \mathbf{A}'\mathbf{G}\mathbf{D}\mathbf{L}_h^T \\ &\quad + \mathbf{L}_i\mathbf{D}\mathbf{G}^T\mathbf{A}'^T + \mathbf{L}_h\mathbf{D}\mathbf{G}^T\mathbf{A}'^T + \mathbf{L}_i\mathbf{D}\mathbf{L}_i^T \\ &\quad + \mathbf{L}_i\mathbf{D}\mathbf{L}_h^T + \mathbf{L}_h\mathbf{D}\mathbf{L}_i^T + \mathbf{L}_h\mathbf{D}\mathbf{L}_h^T\end{aligned} \tag{5.57}$$

by substitution of $\mathbf{G}$ it can be shown that the 2nd and 3rd term of Eqn. (5.57) are zero and the first term corresponds to $\mathbf{A}\mathbf{Q}_{\hat{x}\hat{x}}\mathbf{A}^T$. Writing the last four terms in compact form leads to:

$$\mathbf{K}\mathbf{C}_{zz}\mathbf{K}^T = \mathbf{A}\mathbf{Q}_{\hat{x}\hat{x}}\mathbf{A}^T + (\mathbf{L}_i + \mathbf{L}_h) \cdot \mathbf{D}^{-1} \cdot (\mathbf{L}_i^T + \mathbf{L}_h^T) \tag{5.58}$$

thus, Eqn. (5.54) can be written as:

$$\begin{aligned}\mathbf{Q}_{\hat{l}'\hat{l}'} = {} & \mathbf{C}_{s_i's_i'} + \mathbf{C}_{s_h's_h'} \\ & - (\mathbf{H}_i + \mathbf{H}_h) \cdot (\mathbf{A}\mathbf{Q}_{\hat{x}\hat{x}}\mathbf{A}'^T) - (\mathbf{A}'\mathbf{Q}_{\hat{x}\hat{x}}\mathbf{A}^T) \cdot (\mathbf{H}_i^T + \mathbf{H}_h^T) \\ & - (\mathbf{C}_{s_i's_i} + \mathbf{C}_{s_h's_h}) \cdot (\mathbf{L}_i^T + \mathbf{L}_h^T) - (\mathbf{L}_i + \mathbf{L}_h) \cdot (\mathbf{C}_{s_i's_i}^T + \mathbf{C}_{s_h's_h}^T) \\ & + \mathbf{A}\mathbf{Q}_{\hat{x}\hat{x}}\mathbf{A}^T + (\mathbf{L}_i + \mathbf{L}_h) \cdot \mathbf{D}^{-1} \cdot (\mathbf{L}_i^T + \mathbf{L}_h^T)\end{aligned} \tag{5.59}$$

Using the relation $\mathbf{L} = \mathbf{H} \cdot (\mathbf{I} - \mathbf{A}\mathbf{G}) = \mathbf{H} \cdot \mathbf{R}$ (compare Eqns. (5.45) and (5.46)), matrices $\mathbf{L}_{i,h}^T$ can be written as:

$$\begin{aligned}\mathbf{L}_{i,h}^T &= (\mathbf{I} - \mathbf{G}^T\mathbf{A}^T) \cdot \mathbf{H}_{i,h}^T = (\mathbf{I} - \mathbf{D}^{-1}\mathbf{A}\mathbf{Q}_{\hat{x}\hat{x}}\mathbf{A}^T) \cdot \mathbf{D}^{-1}\mathbf{C}_{s_{i,h}'s_{i,h}}^T \\ &= \mathbf{D}^{-1}(\mathbf{I} - \mathbf{A}\mathbf{Q}_{\hat{x}\hat{x}}\mathbf{A}^T\mathbf{D}^{-1}) \cdot \mathbf{C}_{s_{i,h}'s_{i,h}}^T \\ &= \mathbf{D}^{-1} \cdot \mathbf{R} \cdot \mathbf{C}_{s_{i,h}'s_{i,h}}^T\end{aligned} \tag{5.60}$$

and thus the three terms in Eqn. (5.59) with $\mathbf{L}_i$ and $\mathbf{L}_h$ components can be written as:

$$(\mathbf{L}_i + \mathbf{L}_h) \cdot (\mathbf{C}_{s_i's_i}^T + \mathbf{C}_{s_h's_h}^T) = (\mathbf{H}_i + \mathbf{H}_h)\mathbf{R}(\mathbf{C}_{s_i's_i}^T + \mathbf{C}_{s_h's_h}^T) \tag{5.61}$$

$$\begin{aligned}(\mathbf{C}_{s_i's_i} + \mathbf{C}_{s_h's_h}) \cdot (\mathbf{L}_i^T + \mathbf{L}_h^T) &= (\mathbf{C}_{s_i's_i} + \mathbf{C}_{s_h's_h})\mathbf{D}^{-1}\mathbf{R}(\mathbf{C}_{s_i's_i}^T + \mathbf{C}_{s_h's_h}^T) \\ &= (\mathbf{H}_i + \mathbf{H}_h)\mathbf{R}(\mathbf{C}_{s_i's_i}^T + \mathbf{C}_{s_h's_h}^T)\end{aligned} \tag{5.62}$$

$$\begin{aligned}(\mathbf{L}_i + \mathbf{L}_h)\mathbf{D}(\mathbf{L}_i^T + \mathbf{L}_h^T) &= (\mathbf{H}_i + \mathbf{H}_h)\mathbf{R}\mathbf{D}\mathbf{D}^{-1}\mathbf{R}(\mathbf{C}_{s_i's_i}^T + \mathbf{C}_{s_h's_h}^T) \\ &= (\mathbf{H}_i + \mathbf{H}_h)\mathbf{R}^2(\mathbf{C}_{s_i's_i}^T + \mathbf{C}_{s_h's_h}^T) \\ &= (\mathbf{H}_i + \mathbf{H}_h)\mathbf{R}(\mathbf{C}_{s_i's_i}^T + \mathbf{C}_{s_h's_h}^T)\end{aligned} \tag{5.63}$$

where $\mathbf{R}^2 = \mathbf{R}$ as shown in *Wirth* (1990). By back-substitution of $\mathbf{R}$, the resulting common expresion in Eqn. (5.61), (5.62), and (5.63) can further be expressed as:

$$\begin{aligned}(\mathbf{H}_i + \mathbf{H}_h)\mathbf{R}(\mathbf{C}_{s_i's_i}^T + \mathbf{C}_{s_h's_h}^T) &= (\mathbf{H}_i + \mathbf{H}_h) \cdot (\mathbf{C}_{s_i's_i}^T + \mathbf{C}_{s_h's_h}^T) \\ &\quad - \mathbf{H}_i\mathbf{A}\mathbf{Q}_{\hat{x}\hat{x}}\mathbf{A}^T\mathbf{H}_i^T - \mathbf{H}_i\mathbf{A}\mathbf{Q}_{\hat{x}\hat{x}}\mathbf{A}^T\mathbf{H}_h^T \\ &\quad - \mathbf{H}_h\mathbf{A}\mathbf{Q}_{\hat{x}\hat{x}}\mathbf{A}^T\mathbf{H}_i^T - \mathbf{H}_h\mathbf{A}\mathbf{Q}_{\hat{x}\hat{x}}\mathbf{A}^T\mathbf{H}_h^T \\ &= (\mathbf{H}_i + \mathbf{H}_h) \cdot (\mathbf{C}_{s_i's_i}^T + \mathbf{C}_{s_h's_h}^T) \\ &\quad - (\mathbf{H}_i + \mathbf{H}_h) \cdot \mathbf{A}\mathbf{Q}_{\hat{x}\hat{x}}\mathbf{A}^T(\mathbf{H}_i^T + \mathbf{H}_h^T)\end{aligned} \tag{5.64}$$

Finally, Eqn. (5.59) can be written as:

$$
\begin{aligned}
\mathbf{Q}_{\hat{l}'\hat{l}'} &= \mathbf{C}_{s_i's_i'} + \mathbf{C}_{s_h's_h'} - (\mathbf{H}_i + \mathbf{H}_h) \cdot (\mathbf{A}\mathbf{Q}_{\hat{x}\hat{x}}\mathbf{A}'^T) \\
&\quad + \mathbf{A}\mathbf{Q}_{\hat{x}\hat{x}}\mathbf{A}^T - (\mathbf{A}'\mathbf{Q}_{\hat{x}\hat{x}}\mathbf{A}^T) \cdot (\mathbf{H}_i^T + \mathbf{H}_h^T) \\
&\quad - (\mathbf{H}_i + \mathbf{H}_h) \cdot (\mathbf{C}_{s_i's_i}^T + \mathbf{C}_{s_h's_h}^T) + (\mathbf{H}_i + \mathbf{H}_h) \cdot \mathbf{A}\mathbf{Q}_{\hat{x}\hat{x}}\mathbf{A}^T(\mathbf{H}_i^T + \mathbf{H}_h^T) \\
&= \mathbf{C}_{s_i's_i'} + \mathbf{C}_{s_h's_h'} - (\mathbf{H}_i + \mathbf{H}_h) \cdot (\mathbf{C}_{s_i's_i}^T + \mathbf{C}_{s_h's_h}^T) \\
&\quad \Big[(\mathbf{H}_i + \mathbf{H}_h)\mathbf{A} - \mathbf{A}'\Big]\mathbf{Q}_{\hat{x}\hat{x}}\Big[\mathbf{A}^T(\mathbf{H}_i^T + \mathbf{H}_h^T) - \mathbf{A}'^T\Big]
\end{aligned} \tag{5.65}
$$

Back-substitution of matrix $\mathbf{H}_{i,h} = \mathbf{C}_{s_{i,h}'s_{i,h}}\mathbf{D}^{-1}$ then gives the final expression for the variance-covariance matrix of the total signal derived by the principles of ALSC:

$$
\begin{aligned}
\mathbf{Q}_{\hat{l}'\hat{l}'} &= \mathbf{C}_{s_i's_i'} + \mathbf{C}_{s_h's_h'} \\
&\quad - (\mathbf{C}_{s_i's_i} + \mathbf{C}_{s_h's_h}) \cdot \mathbf{D}^{-1} \cdot (\mathbf{C}_{s_i's_i}^T + \mathbf{C}_{s_h's_h}^T) + \mathbf{M} \cdot \mathbf{Q}_{\hat{x}\hat{x}} \cdot \mathbf{M}^T
\end{aligned} \tag{5.66}
$$

$$
\mathbf{M} = (\mathbf{C}_{s_i's_i} + \mathbf{C}_{s_h's_h}) \cdot \mathbf{D}^{-1} \cdot \mathbf{A} - \mathbf{A}'
$$

$$
\mathbf{Q}_{\hat{x}\hat{x}} = (\mathbf{A}^T\mathbf{D}^{-1}\mathbf{A})^{-1} \tag{5.67}
$$

### 5.4.3 Variance-Covariance of Derived Quantities

In this section, derivation of the variance-covariance matrix of a derived quantity using collocation, like velocity from position (Section 5.1.2), is given here. Using the principle of error propagation presented in Section 5.4.2, the error of a gradient w.r.t. a observation $\mathbf{l}(\mathbf{r})$, $\nabla\mathbf{l}(\mathbf{r})$, in the following written as $\nabla\mathbf{l}$ is expressed as $\epsilon_{\nabla\hat{\mathbf{l}}'} = \nabla\hat{\mathbf{l}}' - \nabla\bar{\mathbf{l}}'$):

$$
\begin{aligned}
\epsilon_{\nabla\hat{\mathbf{l}}'} &= \nabla\hat{\mathbf{l}}' - \nabla\bar{\mathbf{l}}' \\
&= \nabla\mathbf{A}'\hat{\mathbf{x}} + \nabla\hat{\mathbf{s}}' - (\nabla\mathbf{A}'\bar{\mathbf{x}} + \nabla\bar{\mathbf{s}}') \\
&= (\nabla\mathbf{A}'\mathbf{G} + \nabla\mathbf{L})(\mathbf{A}\bar{\mathbf{x}} + \bar{\mathbf{z}}) - (\nabla\mathbf{A}'\bar{\mathbf{x}} + \nabla\bar{\mathbf{s}}') \\
&= (\nabla\mathbf{A}'\mathbf{G}\bar{\mathbf{z}} + \nabla\mathbf{L}\bar{\mathbf{z}}) - \nabla\hat{\mathbf{s}}' \\
&= \nabla\mathbf{K}\bar{\mathbf{z}} - \nabla\hat{\mathbf{s}}'
\end{aligned} \tag{5.68}
$$

as in Eqn. (5.42), with

$$\bar{\mathbf{z}} = \bar{\mathbf{s}} + \bar{\mathbf{n}} \tag{5.69}$$

$$\nabla\mathbf{K} = \nabla\mathbf{A}'\mathbf{G} + \nabla\mathbf{L} \tag{5.70}$$

$$\mathbf{G} = \mathbf{Q}_{\hat{x}\hat{x}}\mathbf{A}^T\mathbf{D}^{-1} \tag{5.71}$$

$$\hat{\mathbf{x}} = \mathbf{G} \cdot \mathbf{l} \tag{5.72}$$

$$\nabla\mathbf{L} = \nabla\mathbf{C}_{s's}\mathbf{D}^{-1} \cdot (\mathbf{I} - \mathbf{A}\mathbf{G}) = \nabla\mathbf{H} \cdot (\mathbf{I} - \mathbf{A}\mathbf{G}) \tag{5.73}$$

$$\nabla\mathbf{L}\mathbf{A} = 0 \tag{5.74}$$

This leads to

$$\epsilon_{\nabla\hat{\mathrm{l}}'}\epsilon^T_{\nabla\hat{\mathrm{l}}'} = (\nabla\mathbf{K}\bar{\mathbf{z}} - \nabla\bar{\mathbf{s}}') \cdot (\bar{\mathbf{z}}^T\nabla\mathbf{K}^T - (\nabla\bar{\mathbf{s}}')^T) \tag{5.75}$$

$$\begin{aligned} &= \nabla\mathbf{K}\bar{\mathbf{z}}\bar{\mathbf{z}}^T\nabla\mathbf{K}^T - \nabla\mathbf{K}\bar{\mathbf{z}}(\nabla\bar{\mathbf{s}}')^T \\ &\quad - \nabla\bar{\mathbf{s}}'\bar{\mathbf{z}}^T\nabla\mathbf{K}^T + \nabla\bar{\mathbf{s}}'(\nabla\bar{\mathbf{s}}')^T \end{aligned} \tag{5.76}$$

Following the strategy presented in Eqns. (5.49) to (5.53), the expectation of the individual components are:

$$E\langle\nabla\bar{\mathbf{s}}'(\nabla\bar{\mathbf{s}}')^T\rangle = \nabla^2\mathbf{C}_{s's'} \tag{5.77}$$

$$E\langle\nabla\bar{\mathbf{s}}'\bar{\mathbf{z}}^T\rangle = \nabla\mathbf{C}_{s's} \tag{5.78}$$

$$E\langle\bar{\mathbf{z}}(\nabla\bar{\mathbf{s}}')^T\rangle = \nabla\mathbf{C}_{ss'} = -\nabla\mathbf{C}_{s's} = \nabla\mathbf{C}^T_{s's} \tag{5.79}$$

$$\begin{aligned} E\langle\epsilon_{\nabla\hat{\mathrm{l}}'}\epsilon^T_{\nabla\hat{\mathrm{l}}'}\rangle &= \mathbf{Q}_{\nabla\hat{\mathrm{l}}'\nabla\hat{\mathrm{l}}'} \\ &= \nabla^2\mathbf{C}_{s's'} - \nabla\mathbf{K}\cdot\nabla\mathbf{C}^T_{s's} - \nabla\mathbf{C}_{s's}\nabla\mathbf{K}^T + \nabla\mathbf{K}\mathbf{C}_{zz}\nabla\mathbf{K}^T \end{aligned} \tag{5.80}$$

Note that the relation of Eqn. 5.79 is only valid if the chosen covariance function is isotrop. The second derivative appearing here is a result of the expectation of a linear combination of cross-covariance functions. Given a covarinace matrix $\mathbf{C}_{ss}(x_a, x_b)$, the covariance matrix $\mathbf{C}_{s's}(x_a, x_b)$ between $s$ and $s'$ in general is given by:

$$\mathbf{C}_{s's}(x_a, x_b) = \mathcal{L}_{s'}|_{x_a}(\mathbf{C}_{ss}(x_a, x_b)) \tag{5.81}$$

with $\mathcal{L}$ being a linear functional (*Guillaume*, 2015). For $\mathbf{C}_{s's'}(x_a, x_b)$ this becomes:

$$\begin{aligned} \mathbf{C}_{s's'}(x_a, x_b) &= \mathcal{L}_{s'}|_{x_b}(\mathbf{C}_{s's}(x_a, x_b)) \\ &= \mathcal{L}_{s'}|_{x_b}\Big(\mathcal{L}_{s'}|_{x_a}(\mathbf{C}_{ss}(x_a, x_b))\Big) \\ &= \mathcal{L}_{s'}|_{x_a}\Big(\mathcal{L}_{s'}|_{x_b}(\mathbf{C}_{ss}(x_a, x_b))\Big) \end{aligned} \tag{5.82}$$

In the derivation here, $\mathcal{L}$ is the derivative with respect to $x_a$ and $x_b$.

The final expression for the error matrix of the strain is then obtained by back-substitution of $\nabla\mathbf{K}$, following the same rearrangement principles as in the previous section. Eqn. 5.80 can then be written as:

$$\mathbf{Q}_{\nabla\hat{\mathrm{l}}'\nabla\hat{\mathrm{l}}'} = \nabla^2\mathbf{C}_{s's'} - \nabla\mathbf{C}_{s's}\mathbf{D}^{-1}\nabla\mathbf{C}^T_{s's} + \nabla\mathbf{M}\cdot\mathbf{Q}_{xx}\nabla\mathbf{M}^T \tag{5.83a}$$

$$\nabla\mathbf{M} = \nabla\mathbf{C}_{s's}\mathbf{D}^{-1}\mathbf{A} - \nabla\mathbf{A}' \tag{5.83b}$$

Although not used in this study, error propagation for the strain field introduced in Section 5.3.4 for the ALSC process, can be derived as well. More details therefore are given in the Appendix D.1.

# 6 Results

Following the theoretical derivations and simulations given in the previous chapters, the results for the estimated displacements of the southern part of the Grabengufer sagging area are presented here. Section 6.1 first presents and describes the results obtained when a single camera, combined with a DEM, is used. In Section 6.2, the principle of bundle adjustment is used to estimate object coordinates for the various measurement epochs. Regarding collocation, Section 6.3 shows the applied collocation technique used also to estimate respective velocities in the former two sections. A short presentation of results obtained by applying the technique of ALSC is finally given in Section 6.4.

## 6.1 Single-View Velocity Estimation

In Section 4.2, the topic of estimating metric displacements using a single camera combined with a DEM was addressed. Although two cameras were used in this study, a comparison in terms of estimated velocities is given here. The average direction of motion is nearly Northwest (azimuth between 310 and 330 degrees), thus the projected motion is larger for camera station 1 that is used for this comparison.

### 6.1.1 Correction of Camera Motion

When building a time series of features moving in the image space, the first task is to correct for any type of camera motion. In this process, image coordinates are aligned to a reference epoch (reference orientation for $t = t_0$) for which, e.g., a projection onto a DEM is carried out. With respect to such a reference epoch, image coordinate positions are affected by the variation of camera parameters (extrinsic, intrinsic, and additional parameters). Using a single camera, the principles of the camera parameter estimation procedure (a subproblem of bundle adjustment, Section 4.4.7 and C.2) can be used. Optimally, stable areas or points of known position (ground control points) are known around the object of interest. As discussed in Section 4.4.7, a good network is required to reliably solve for all the camera parameters. For the present geometry, areas of solid rock project into a diagonally shaped region in the upper half of the image (see Fig. 6.6). Such a geometry does not allow for a reliable estimation of camera parameters necessary to correct all the main components of apparent camera motion. E.g., high correlations are observed between the first radial distortion parameter $k_1$ (Eqn. (4.15)) and the camera constant $c$ (Eqn. (4.14)), as well as for the principal point offset $y_p$ and the rotation angle $\omega$, all varying over time.

Using a simplified model, as discussed in Section 4.2.3, with constraints on high-order distortion parameters was found to be not a good choice: approximating the projection of near- and far-field objects seen by camera 1, led to large residuals, specially along the right image edge with objects being relatively close.

To allow a comparison, camera parameters estimated during the bundle adjustment (Section 6.2.5) are used to rectify image coordinates with respect to the corrected camera coordinates of epoch $t = t_0$. The strategy presented in Section 4.2.3, Eqn. (4.12) is used for this purpose.

Including the additional parameters to compute the corrected image coordinates $(..)_{\text{corr}}$, a total of $2 \cdot 16 = 32$ parameters are involved. Given the errors of both components, error propagation for the referenced image coordinates $x_{\text{ref}}$ and $y_{\text{ref}}$ yield higher values than the estimated matching errors for image coordinates used in the bundle adjustment.

An estimation of significant parameters using a single view in the current geometry led to errors in the order of 3 and 8 pixels (an example is shown in Section 6.2.5 for the estimated template image coordinates). Using the estimated camera parameters determined during the bundle adjustment shows a considerable improvement resulting in errors in the order of 0.5 to 1.5 pixels, generally increasing toward the image corners (as errors in distortion terms become more significant). This shows that the constrained geometry using at least two views not only allows for a direct reconstruction but also leads to camera parameters of higher precision.

### 6.1.2 Velocity Estimates

Time series of features tracked and corrected in the image space are being filtered using the method of collocation (see Section 6.3). Empirical covariance functions (as described in Section 5.2) were estimated in a series of feature time sequences. With $d$ being the time lag, the following function with parameters $u$, $n$, and $r_s$ was found to be well suited in average:

$$
\begin{aligned}
f(d) &= \sigma_s^2 \cdot \frac{1}{1 + u \cdot |d|^n} \qquad (6.1)\\
\text{with} \quad \sigma_s^2 &= 1 \ [pixel^2]\\
n &= 2\\
r_s &= 120 \ [\text{days}] \quad (u = {}^1\!/_{r_s^n} \sim 7 \cdot 10^{-5})
\end{aligned}
$$

Using Eqn. (6.1) to define the covariance matrix, velocities and the respective errors were estimated as described by Eqns. (5.6) and (5.83). Being still in the image space, velocities are given in units of pixel/day. A time series of rectified image coordinates, both as unfiltered and filtered (i.e., collocated) estimates, is shown in Fig. 6.1.

Following the principles described in Section 4.2.1, ray tracing is applied to convert the velocities in the image space into a metric unit. The DEM (provided by the Federal Office

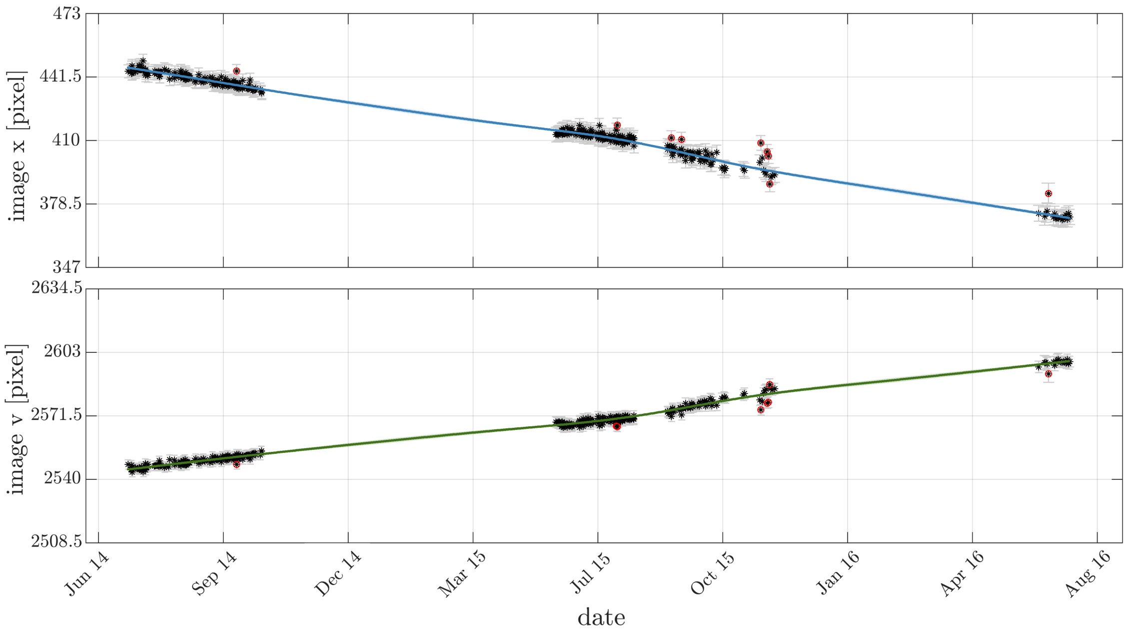


**Figure 6.1:** Collocated time series of a rectified feature trajectory in the image space. The principle coordinate estimates were obtained by least-squares feature matching. Coordinate corrections for each epoch were conducted using the camera parameters estimated during the bundle adjustment. Position estimates encircled in red are outliers, defined as positions with a large deviation w.r.t. its collocated results.

of Environment) was used, whereas an standard deviation of 1 m in the height component was assigned for the DEM. As described in Section 4.2.1, errors in a DEM heavily depend on the method it was generated with and are not uniformly distributed (e.g., correlated with slope steepness). As shown in Section 4.2.2, the validity of formal error propagation also heavily depends on the geometry, i.e., the projection angle, and is favorably determined by a Monte Carlo simulation.

Whenever an intersection point on the DEM is found, its distance $t$ (compare Eqn. (4.5)) from the camera center can be computed. Given the pixel size and the focal length of the camera, the velocity scaling $v_s$ is given by:

$$v_s = \frac{\text{pixel size}}{\text{focal length}} \cdot t = s \cdot t \tag{6.2}$$

The pixel size of the camera (Nikon D300s) is 5.5 $\mu$m and the focal length as determined during the bundle adjustment (i.e., around 13.8 mm). For the purpose of error propagation, the scaling coefficient $s$ (Eqn. (6.2)) is included in the full variance-covariance matrix of all observations that are involved.

Fig. 6.2 shows an example of scaled velocity estimates with errors given for each component (i.e., $v_x$ and $v_y$) for June 2014. A black ellipse is drawn around the area of critical error propagation: as shown by the Monte Carlo simulation in Section 4.2.2, the formal

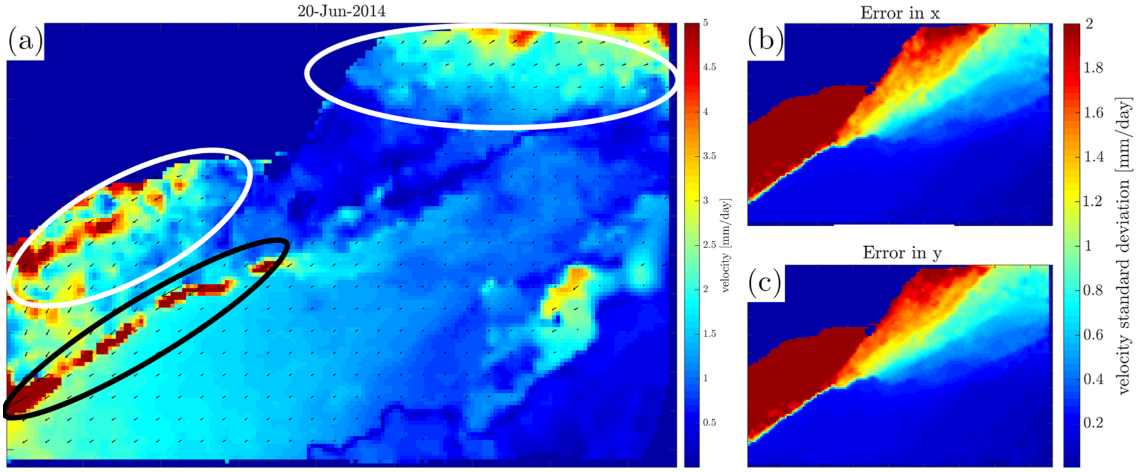


**Figure 6.2:** (a) Absolute velocity estimates (2D) obtained by projection onto a DEM. Encircled in black shows the critical projection area, areas indicated by white ellipses show heavily amplified rectification residuals. (b) and (c) show the vertical and horizontal errors of absolute velocities (given by error propagation), respectively.

error is heavily underestimated in this region. The DEM used for this projection was taken in 2010, i.e., four years earlier. Specially in areas close to the camera, the accuracy of the DEM is poor as the rock agglomeration has moved farther downstream. As a result, most of the DEM heights in this area are too low and scaling values derived with DEM intersection points for corresponding rays become overestimated. The two effects (DEM error and geometrical aspects of error propagation for this geometry) lead to velocity estimates that are heavily overestimated (red areas in the highlighted zone of Fig. 6.2a).

Also visible in Fig. 6.2 are (originally) small residuals of radial distortion inaccuracies (highlighted in white). Although not significant (see Fig. 6.3), large scaling values for distant intersection points amplify these velocities to a high degree. In the hypothetical case of a wrong radial distortion parameter (yielding radially symmetric displacements), the result seen in Fig. 6.3 would not be radially symmetric because different scaling values apply to the various regions as they are at different distances. Thus a systematic error of a single radial distortion parameter cannot be responsible for the observed effect.

Combining the two results of Fig. 6.2, i.e., showing only significant velocities, a more realistic picture of the scene is obtained. The significance level is defined by the magnitude of the velocity being larger than its $3\,\sigma$ level. A sequence of estimated velocities above this significance level is shown in Fig. 6.3 for the summer months of 2014 with a time interval of 20 days. To increase the accuracy in the near field area, the DEM was corrected using the results obtained with the bundle adjustment (Section 6.2.5). The sequence shows a general acceleration towards the late summer months (September) before it decelerates again. In the right part of the scene visible in this view, an isolated area with high acceleration (from about $2\,\mathrm{mm/day}$ to $5\,\mathrm{mm/day}$) can be observed. For reference, the same sequence showing also the non-significant areas is given in Fig. 6.4.

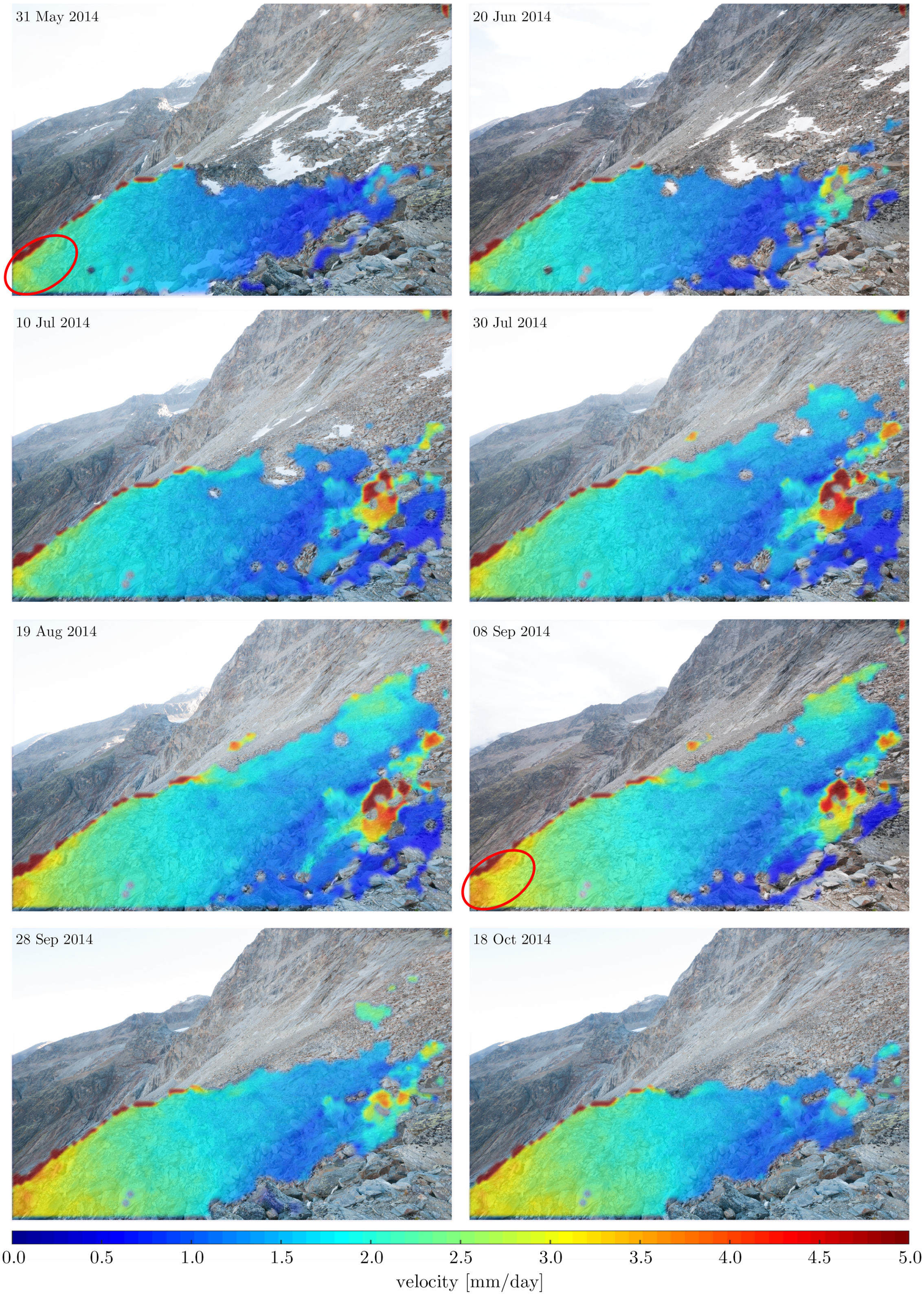


**Figure 6.3:** Series of absolute velocity estimates (2D) showing a small acceleration during the summer months of 2014, i.e., the front area (red ellipse) between September and May shows a velocity increase of about 40 %. The time separation between presented results is 20 days. Only significant velocities ($v \leq 3 \cdot \sigma_v$) are shown.

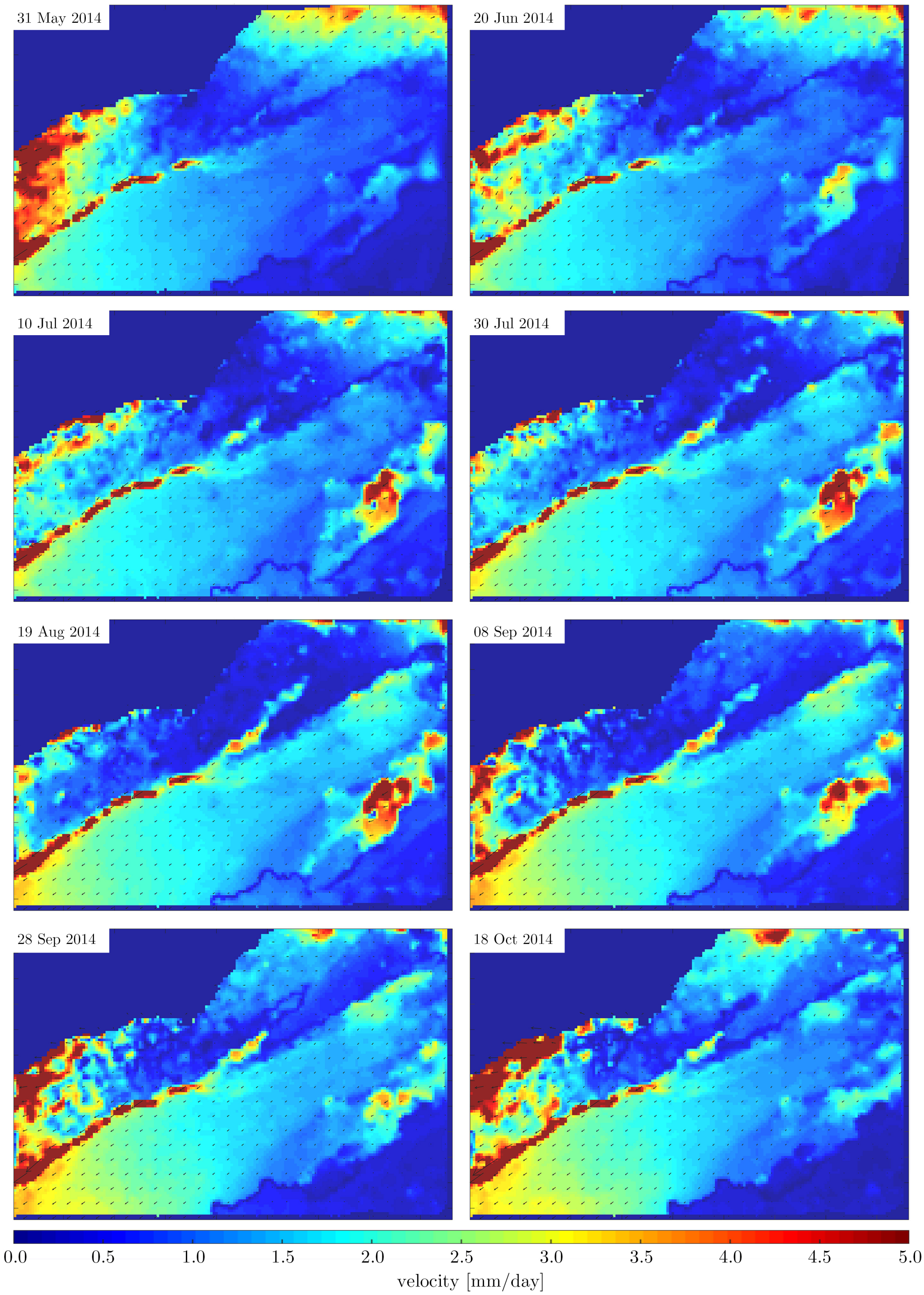


**Figure 6.4:** The same sequence as in Fig. 6.3 but showing also the non-significant flow velocities.

## 6.2 Object Point Reconstruction

In this section, results of object point reconstructions using the principle of bundle adjustment is presented. A schematic view of the near-field geometry is shown in Fig. 6.5. With a baseline of 42.3 m between the two camera centers, near-field objects (with 50 m mean distance) are seen with an angular view difference of more than 56 degrees. Because the surface topography of the observed area is exceptionally rough, the differences in the projection angles cause difficulties for feature matching between the two views. Thus an initial set of well defined features[1] seen in both views, was manually identified. For half of these features, successful matching between the views was possible, yielding accurate point correspondences in the sub-pixel range. Fig. 6.6 shows the distribution and type of points that were placed in both views: in blue are features, where significant displacement is expected, in green are points on solid rock (stable), and squares in red show the image location of the two GNSS stations. Black triangles represent SURF[2] feature correspondences (see also Section 3.2.5). These are determined independently between the views for each epoch and are used to increase robustness and redundancy during the bundle adjustment. The various features defined for tracking are labeled in Fig. 6.7 and Fig. 6.7, respectively. For all theses points, tracking was performed for the complete image

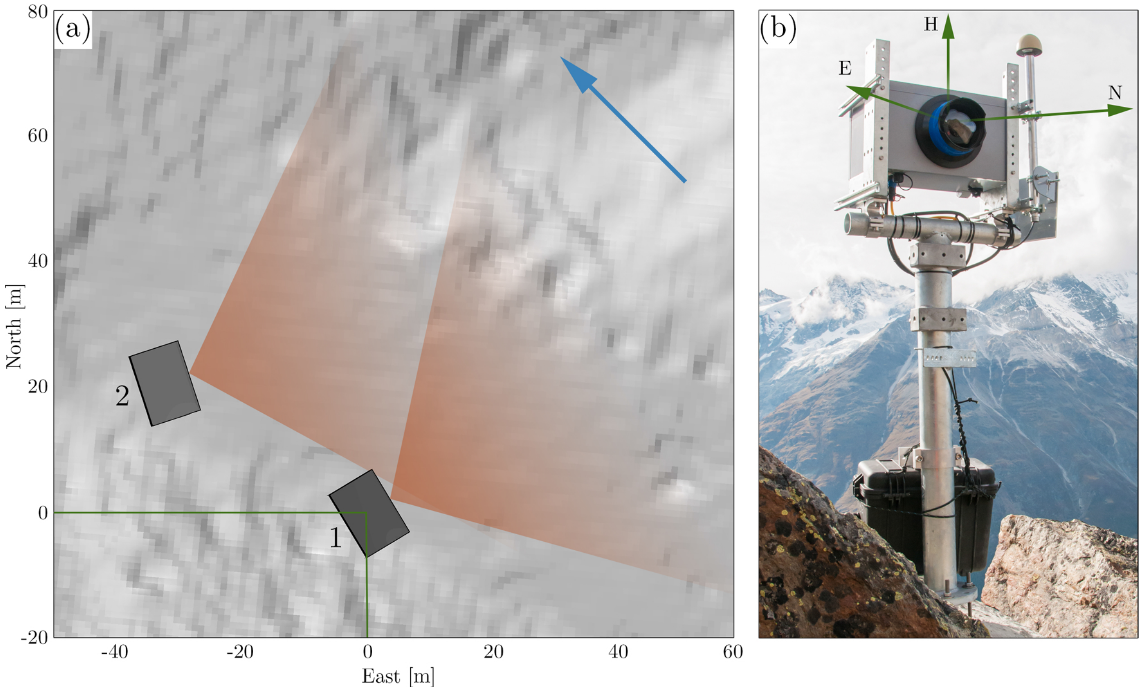


**Figure 6.5:** (a) Schematic view and orientation of the two cameras next to the observation area. The image background shows a shaded DEM relief provided by the Federal Office of Environment. The blue arrow indicates the general flow direction of the permafrost creep. (b) Station 1 with East-, North, and Height axis (in green) showing the topocenter of the local coordinate system. The antenna is at an offset of roughly 50 cm from the origin of the coordinate system.

[1] in terms of high contrast and uniqueness
[2] Speed-Up Robust Features

sequence spanning almost four years. Due to the slow movement of this permafrost area ($\sim 60\,{}^{\mathrm{cm}}/_{\mathrm{year}}$), 92% of them were successfully tracked over this period.

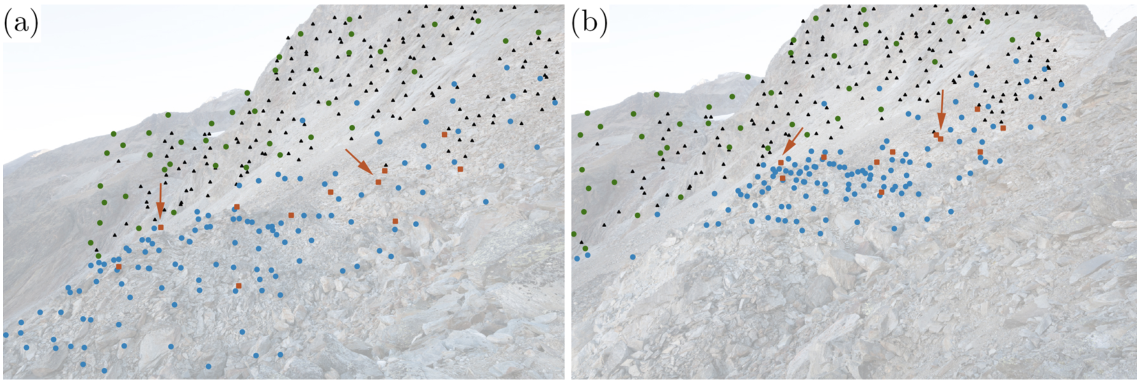


**Figure 6.6:** Measurement points used for object point reconstruction for (a) view 1 and (b) view 2. Positions in blue show features that are expected to move, points in green are those found on solid rock, and the GNSS stations used for calibration are shown in red. Arrows point to the permanent GNSS stations. Along with points on solid rock, the permanent GNSS stations are used as ground control points for all epochs $t > t_0$, with $t_0$ being the initial calibration epoch. Black features are SURF descriptors used as additional features seen in both views.

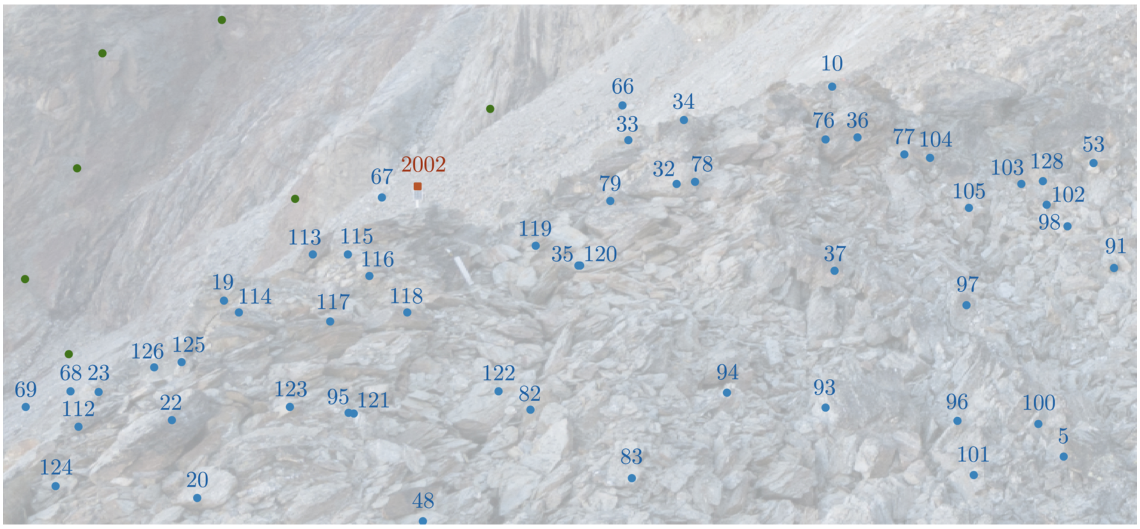


**Figure 6.7:** Detailed view of *ID* numbers associated to individual rock boulders. The complete view is shown in Fig. 6.8.

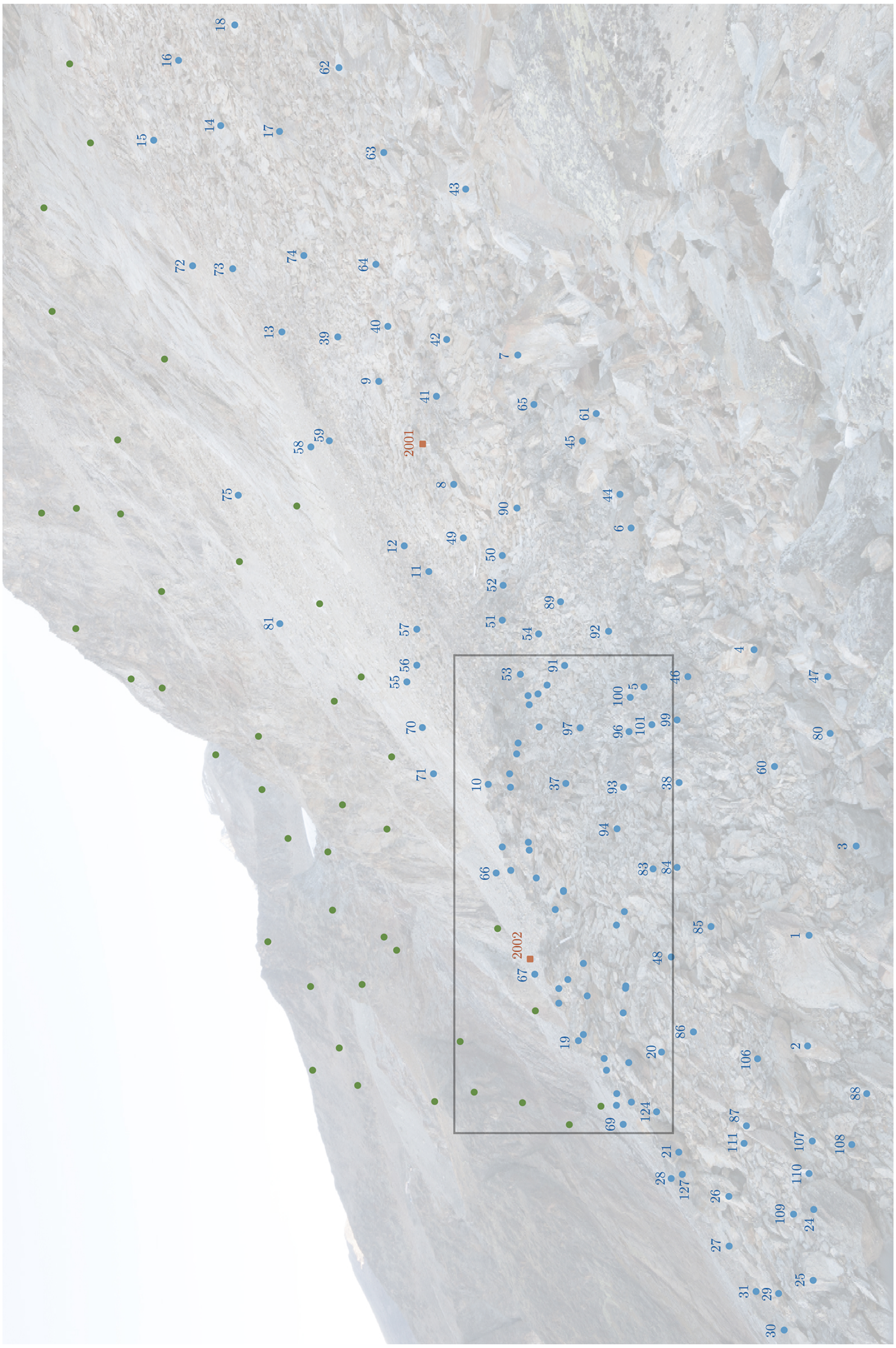

**Figure 6.8:** *ID* numbers associated with the results presented in this section. Note that the numbering has no specific order nor is it related to specific areas of the moving area. The rectangular area is shown in more detail in Fig. 6.7.

### 6.2.1 Initial Solution and Accuracy of Reconstruction

An initial estimation of the various object point coordinates was performed using the calibration patterns (Section 4.5.4) as well as the two GNSS stations (antenna coordinates) as ground control points. The computation was carried out with large weights on the additional parameters because these were determined by a separate calibration procedure (Australis software package, (*Photometrix*, 2013)) using the calibration board (Fig. 4.10).

With 11 ground control points for the initial epoch $t_0$, the quality of absolute coordinate reproduction is evaluated by using only 10 of them at a time. Table 6.1 shows the coordinate differences obtained, when each of the ground control points was once estimated rather than being included as known position. The first two *IDs* correspond to the GNSS stations, whereas the other nine positions are the calibration patterns (Fig. 4.13). According to Eqn. (4.42), the hypothesis test suggested that ground control point $ID = 03$, is erroneous, thus it was rejected for the initial solution. The East component shows to be the worst in terms of precision, a result of the geometrical orientation of the two cameras (i.e., the prominent heading is in East direction).

The result of the estimated accuracy (following *Baarda* (1967), i.e., the precision of the estimated parameters as well as their external reliability) is shown in Fig. 6.9. In the close range area ($50 - 200$ m) the parameter standard deviations are in the order of $2 - 15$ cm, continuously increasing with increasing distance. For this geometry, the Height component is the most accurate, followed by the North and East components. The latter shows stronger variations because of the geometrical orientation of the cameras with respect to the observed area (i.e., near-field points are oriented mostly in East-West direction). The

**Table 6.1:** Evaluation of absolute coordinate estimation accuracy. Each *ID* corresponds to a different ground control point at the initial calibration stage (the first two being the two permanent GNSS stations). Results show the difference between coordinates derived by GNSS and those derived by photogrammetric bundle adjustment. The second column gives the average image sampling from both views.

| *ID* | sampling [cm/pixel] | $\Delta E$ | $\Delta N$ [cm] | $\Delta H$ | $\hat{\sigma}_E$ | $\hat{\sigma}_N$ [cm] | $\hat{\sigma}_H$ |
|---|---|---|---|---|---|---|---|
| 01 | 5.7 | -20.1 | 0.3 | -3.9 | 3.3 | 0.8 | 0.9 |
| 02 | 4.2 | 5.1 | 12.6 | -2.6 | 2.5 | 1.2 | 1.5 |
| 03 | 6.5 | 23.3 | 0.4 | 23.8 | 8.7 | 3.2 | 1.2 |
| 04 | 3.3 | 0.2 | -10.5 | -3.3 | 2.3 | 2.1 | 0.9 |
| 05 | 5.3 | -13.0 | -10.9 | -29.3 | 6.0 | 3.9 | 1.1 |
| 06 | 5.0 | 9.6 | 4.6 | -20.1 | 5.3 | 2.5 | 0.9 |
| 07 | 3.2 | -8.4 | -5.3 | 3.8 | 2.4 | 1.5 | 1.0 |
| 08 | 3.9 | 11.4 | -6.2 | 2.9 | 4.2 | 1.2 | 1.1 |
| 09 | 6.7 | 4.2 | 1.8 | -31.3 | 11.2 | 2.3 | 1.4 |
| 10 | 5.8 | 43.8 | -10.7 | -6.5 | 10.0 | 1.5 | 1.3 |
| 11 | 7.9 | 33.6 | 3.0 | -27.8 | 18.9 | 1.9 | 3.4 |

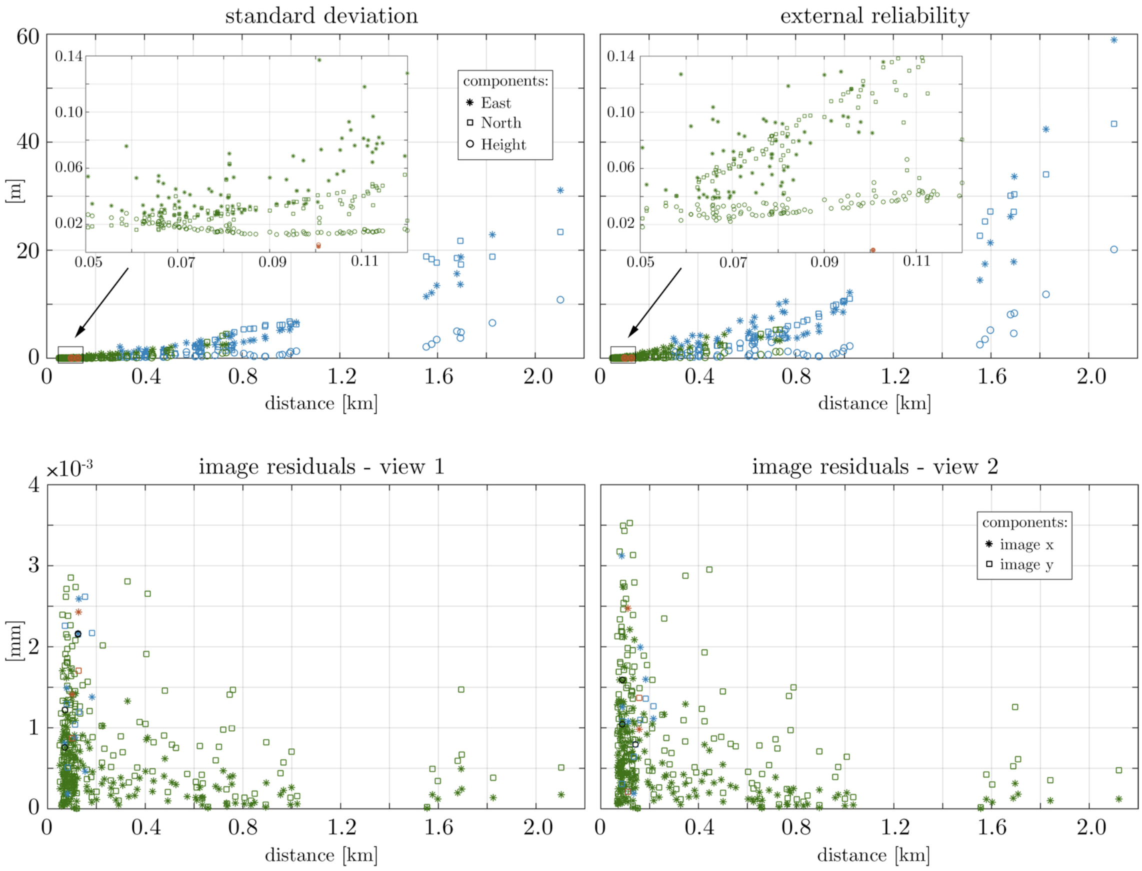


**Figure 6.9:** Accuracy graphs of initial bundle adjustment, i.e., using the absolute calibration patterns (Fig. 4.13) and the GNSS stations as ground control points. The top row illustrates the effect of absolute errors (standard deviations) as well as the external reliability of the reconstructed object points. Insets show the achieved accuracy for the near field. The bottom row shows the corresponding image residuals for both components. Color coding is equivalent to the points marked in Fig. 6.6. Image residuals for the near field are in the order of $2 \times 10^{-3}$ mm (or $\sim$ 0.4 pixel).

external reliability generally shows larger values, where the North component for the near-field is equally affected by undetected gross errors as the East component.

Considering image coordinate residuals (bottom row in Fig. 6.9), the following observations can be made: objects close to the cameras are better determined due to their geometrical position (compare also Fig. 6.10) and thus more constrained. As a result, discrepancies of intersecting rays are more easily seen as image residuals. For distant objects, the direction parallel to the line-of-sight is weakly bounded, and thus image residuals parallel to the epipolar plane become smaller (also discussed in Section 4.4.7), resulting in a general decrease of image residuals of distant objects. In addition, close-range features could not be matched by LSM between the views (only manual identification), also contributing to larger image residuals.

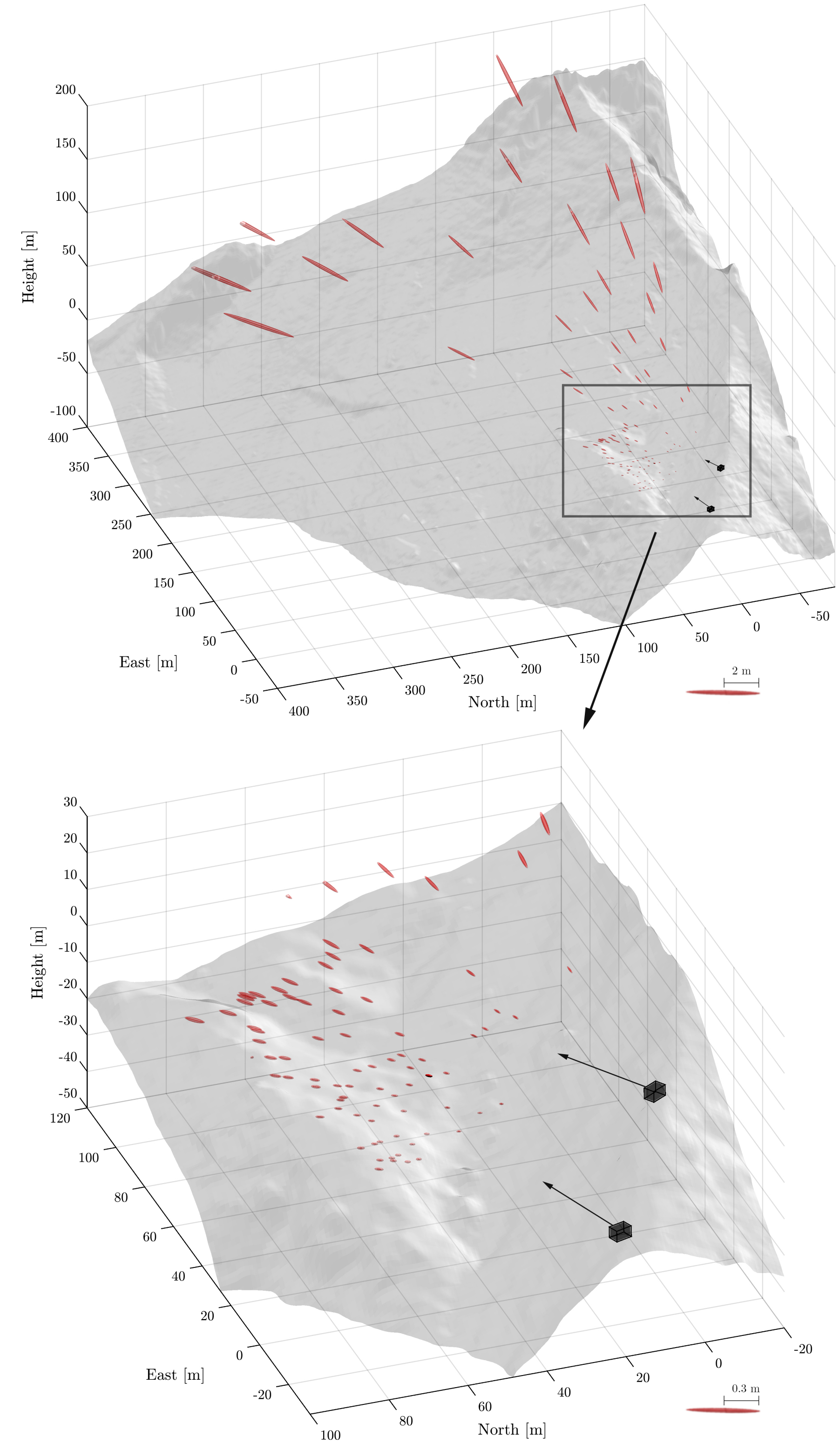


**Figure 6.10:** 95% error ellipsoids derived during the initial bundle adjustment at epoch $t = t_0$. DEM provided by the Federal Office of the Environment (FOEN).

Fig. 6.10 shows the distribution and orientation of the 95% error ellipsoids obtained for the initial solution of the bundle adjustment. The highest variance components are in viewing direction of the cameras, that is mostly East-Northeast. Pseudo-control points used for epochs $t > t_0$ exhibit the largest extensions as they are located on relatively distant solid rock.

The offset between the camera centers and the GNSS centers were determined in the initial adjustment procedure. For this purpose the camera position parameters were treated as free parameters (i.e., without including them as pseudo-observations). With respect to the GNSS centers, the following offsets were determined:

- Camera 1: $\Delta_{E,N,H} = 4.4, -35.9, -38.5\,\text{cm}$
- Camera 2: $\Delta_{E,N,H} = 15.5, -28.3, -38.4\,\text{cm}$

with standard deviations in the order of $\sigma_E = 6\,\text{cm}$ and $\sigma_{N,H} = 2.5$ cm for both cameras. These offset values agree well with the expected offset as shown in Fig. 6.5. Over the course of the study period, the absolute positions of the cameras changed by a few centimeters (Appendix E).

### 6.2.2 Determinability of Camera Parameters

Section 4.4.7 showed that a strong network is required to estimate additional parameters with adequate significance. In this section, the determinability of intrinsic and additional parameters, with respect to the given camera installation geometry, for object point reconstruction is presented. In Fig. 6.6, the geometrical point distribution with respect to both views is shown. Especially for station 2, common object points occupy only about half the image space. As this condition might lead to problems for estimating significant intrinsic and additional parameters, a synthetic test was performed: the initial set of object points are used to generate synthetic image coordinates in both views. Image coordinates are then disturbed by Gaussian noise $\mathbf{n} \sim \mathcal{N}(0, \sigma_{1,2,3})$, whereas $\sigma_1$, $\sigma_2$, and $\sigma_3$ correspond to different noise levels in pixel units. For the ground control point accuracies, estimated errors obtained with the initial reconstruction process were used (blue markers in Fig. 6.9).

Table 6.2 shows that even for low noise levels, not all the additional parameters can be determined with significance for both, camera 1 and camera 2, respectively. Given the initial reconstruction result (Section 6.2.1), the average noise level is in the order of 0.4 pixel. As the results of this simulation suggests, the additional parameters $k_1$, $k_2$, $k_3$, $p_1$, and $p_2$ as well as the principal point position $x_p$ and $y_p$ are treated as pseudo observations in the adjustment for epochs $t > t_0$. High order parameters were completely fixed by setting large weights on the respective pseudo-observations while other parameters where given more freedom for adjustment (more details in Section 6.2.5).

The accuracy of the photogrammetrically estimated camera positions are in the order of the measured camera motion over the measurement period (Appendix E), thus these parameters are also treated as pseudo-observations with large weights (see next Section).

**Table 6.2:** Results of camera parameter reconstruction for camera 1 and camera 2 using synthetic image coordinate data and the principles of bundle adjustment. For each image noise level $\sigma_i$, all parameters are estimated without adding pre-known parameters as pseudo-observations. The first columns, $\hat{p}$, contain the estimated parameters whereas the second columns, $\hat{\sigma}_p$, show the respective parameter standard deviations. Fields marked by $'-'$, were found not to be significant during the adjustment, thus constrained at zero. For $O_X$, $O_Y$, and $O_Z$, units are in meter, for $\omega$, $\phi$, and $\kappa$ in degrees, and for all other parameters in the metric camera units (mm). The last row, $\hat{\sigma}_0$, lists the respective a posterior standard deviation.

| Cam 1 | $\check{p}$ | $\sigma_i = 0.200$ | | $\sigma_i = 0.600$ | | $\sigma_i = 1.500$ | |
|---|---|---|---|---|---|---|---|
| | | $\hat{p}$ | $\hat{\sigma}_p$ | $\hat{p}$ | $\hat{\sigma}_p$ | $\hat{p}$ | $\hat{\sigma}_p$ |
| $x_p$ | 0.158 | 0.121 | 0.018 | - | - | 0.108 | 0.034 |
| $y_p$ | $-0.064$ | $-0.094$ | 0.015 | 0.081 | 0.018 | - | - |
| $c$ | 13.716 | 13.701 | 0.014 | 13.728 | 0.016 | 13.712 | 0.023 |
| $k_1$ | $-1.35e^{-4}$ | $-1.50e^{-4}$ | $0.05e^{-4}$ | $-1.10e^{-4}$ | $0.08e^{-4}$ | $-1.09e^{-4}$ | $0.18e^{-4}$ |
| $k_2$ | $6.82e^{-7}$ | $9.30e^{-7}$ | $0.32e^{-7}$ | - | - | - | - |
| $k_3$ | $1.02e^{-9}$ | - | - | $4.58e^{-9}$ | $0.29e^{-9}$ | $4.03e^{-9}$ | $0.59e^{-9}$ |
| $p_1$ | $-5.36e^{-5}$ | - | - | - | - | - | - |
| $p_2$ | $-8.02e^{-5}$ | $-4.29e^{-5}$ | $0.90e^{-5}$ | $-13.01e^{-5}$ | $2.31e^{-5}$ | $-12.07e^{-5}$ | $2.80e^{-5}$ |
| $\omega$ | 4.657 | 4.653 | 0.002 | 4.678 | 0.003 | 4.665 | 0.000 |
| $\phi$ | 4.169 | 4.167 | 0.001 | 4.159 | 0.000 | 4.165 | 0.002 |
| $\kappa$ | 3.089 | 3.084 | 0.002 | 3.104 | 0.002 | 3.092 | 0.001 |
| $O_X$ | 0.108 | 0.205 | 0.099 | 0.020 | 0.117 | 0.226 | 0.167 |
| $O_Y$ | $-0.193$ | $-0.089$ | 0.047 | $-0.261$ | 0.078 | $-0.138$ | 0.125 |
| $O_Z$ | $-0.072$ | $-0.066$ | 0.016 | 0.030 | 0.024 | 0.024 | 0.046 |
| **Cam 2** | $\check{p}$ | $\sigma_i = 0.200$ | | $\sigma_i = 0.600$ | | $\sigma_i = 1.500$ | |
| | | $\hat{p}$ | $\hat{\sigma}_p$ | $\hat{p}$ | $\hat{\sigma}_p$ | $\hat{p}$ | $\hat{\sigma}_p$ |
| $x_p$ | 0.237 | 0.201 | 0.022 | 0.219 | 0.022 | 0.224 | 0.032 |
| $y_p$ | $-0.062$ | $-0.084$ | 0.012 | - | - | - | - |
| $c$ | 13.876 | 13.889 | 0.015 | 13.910 | 0.016 | 13.854 | 0.023 |
| $k_1$ | $-1.30e^{-4}$ | $-1.14e^{-4}$ | $0.04e^{-4}$ | $-0.89e^{-4}$ | $0.07e^{-4}$ | $-1.26e^{-4}$ | $0.18e^{-4}$ |
| $k_2$ | $6.61e^{-7}$ | $5.96e^{-7}$ | $0.26e^{-7}$ | - | - | - | - |
| $k_3$ | $2.46e^{-9}$ | - | - | $2.78e^{-9}$ | $0.27e^{-9}$ | $4.44e^{-9}$ | $0.65e^{-9}$ |
| $p_1$ | $2.94e^{-5}$ | $7.28e^{-5}$ | $2.00e^{-5}$ | - | - | - | - |
| $p_2$ | $5.10e^{-5}$ | $8.59e^{-5}$ | $1.08e^{-5}$ | $-8.04e^{-5}$ | $1.10e^{-5}$ | - | - |
| $\omega$ | 4.556 | 4.553 | 0.003 | 4.571 | 0.001 | 4.568 | 0.002 |
| $\phi$ | 4.391 | 4.389 | 0.001 | 4.392 | 0.001 | 4.392 | 0.002 |
| $\kappa$ | 2.992 | 2.989 | 0.003 | 3.006 | 0.001 | 3.002 | 0.002 |
| $O_X$ | $-32.712$ | $-32.803$ | 0.110 | $-32.855$ | 0.132 | $-32.387$ | 0.246 |
| $O_Y$ | 20.450 | 20.446 | 0.040 | 20.450 | 0.062 | 20.600 | 0.119 |
| $O_Z$ | $-16.847$ | $-16.863$ | 0.018 | $-16.845$ | 0.029 | $-16.719$ | 0.058 |
| $\hat{\sigma}_0$ | - | 0.155 | | 0.509 | | 1.265 | |

### 6.2.3 Effect of GNSS Integration

What so far has not been addressed is the question about the effect of the GNSS solutions with respect to the accuracy of object point reconstruction. As proposed in Section 6.2.2, initially determined coordinates of object points located on solid rock are used as pseudo ground control points for the reconstruction process in the forthcoming epochs. With these, the condition for a successful estimation of camera parameters and new object point coordinates in its principle is given.

Accuracy estimates without including any GNSS station measurements are shown in the top row of Fig. 6.11: the reconstruction precision for the near-field features is in the order of 40 cm for the North and Height and around 60 cm for the East component. Adding GNSS station 1 in the field of view (FOV) improves the coordinate precision by a factor of 3, the external reliability, however, remains nearly unchanged. A different result is obtained if GNSS station 2 instead of 1 is included. Here, the precision of the Height component is below the 5 cm level and just above the 5 cm level for the East and North components. A similar improvement is found for the external reliability. Using both GNSS stations as ground control points further improves the precision for all components. The last row in Fig. 6.11 also shows the effect of adding an additional GNSS station closer to the cameras: the most prominent outcome here is seen in the improved external reliability for near-field objects.

The most severe effect of including GNSS solutions into the bundle adjustment is obtained when the camera positions are being constraint (see Fig. 6.12). As the origin of all bundles with respect to the coordinate system become fixed, the overall stability is improved. This also shows that external reliability is strongly bounded to the precision of known camera positions. Adding GNSS stations as ground control points in the field of view further improves the accuracy (Fig. 6.12). For the current geometry, one additional station placed at a distance between 50 and 100 m improves the precision by about a factor of two for points being also at such distances from the cameras. This effect is obtained when GNSS station 2, being at an average distance of 101 m, is included. For station 1, no significant improvement is observed. Care however must be taken in this analysis, because the results also depend on the image coordinate accuracies as well as on the number of points used for reconstruction. For example, increasing the number of image point correspondences between the views increases the redundancy, thus higher accuracies can be expected.

For the time series presented in Section 6.2.5, camera position parameters were constrained to the GNSS position estimates of the respective station for every epoch. These coordinates were corrected by the offset determined during the initial adjustment procedure (Section 6.2.1). Both camera displacements are in the order of a few centimeters over the course of the study period (see Appendix E).

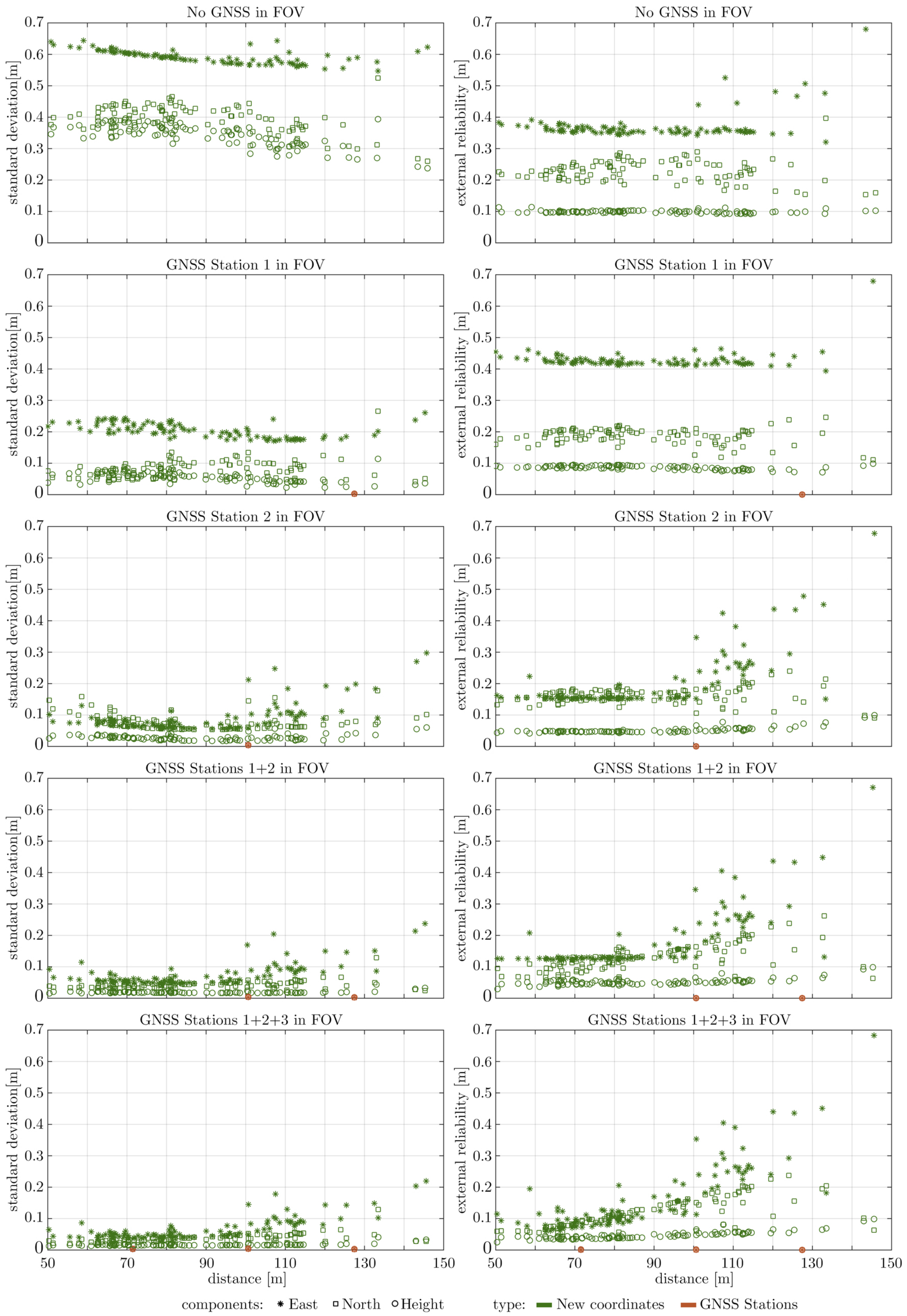


**Figure 6.11:** Effect of adding GNSS stations in the field of view (FOV) of both cameras. These results were computed by estimating all camera parameters as unconstrained variables. The left column shows the estimated parameter standard deviations as a function of mean distance. In the right column the corresponding external reliability estimates are plotted.

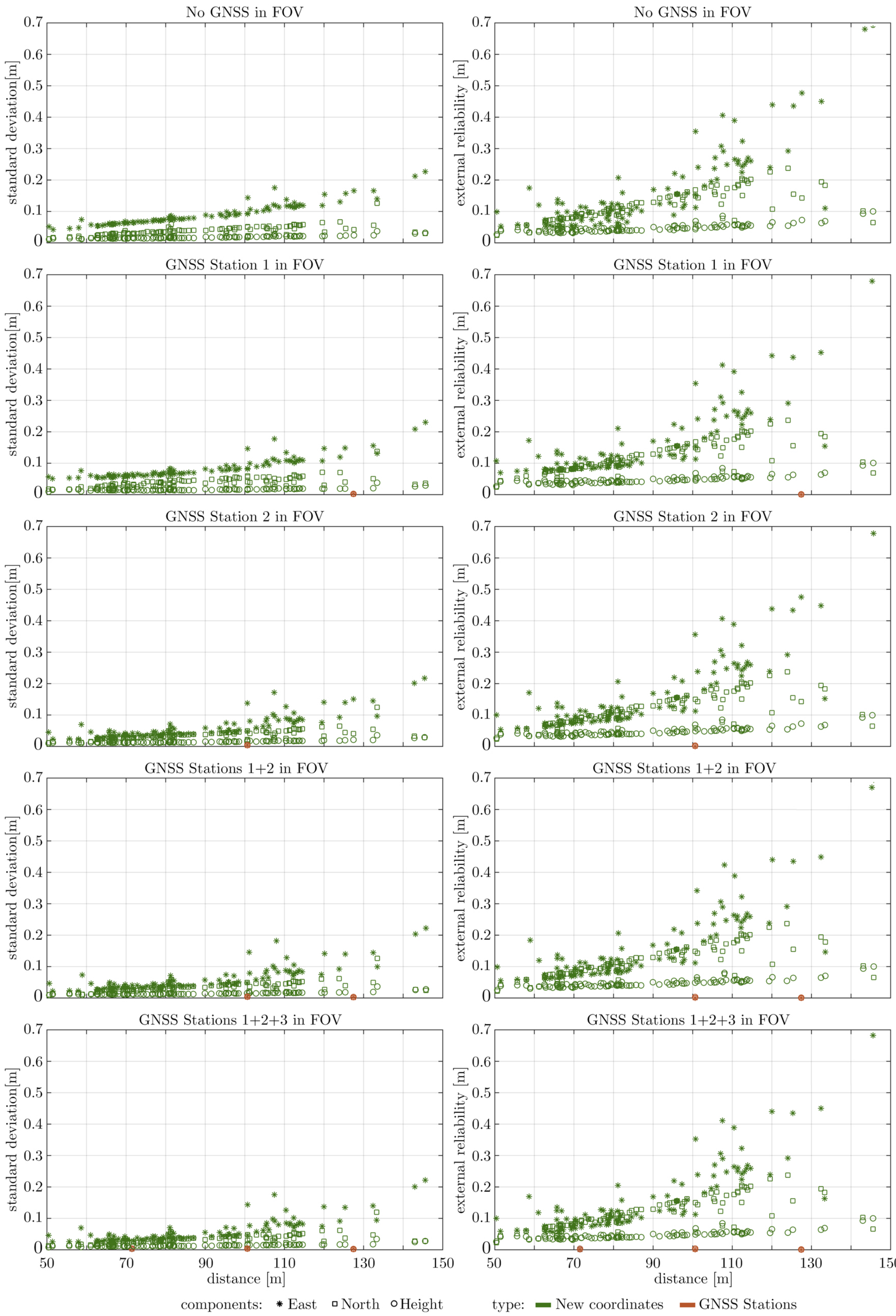


**Figure 6.12:** Effect of adding GNSS stations in the field of view (FOV) of both cameras. With respect to Fig. 6.11, the difference here is that both camera positions were constrained by the respective GNSS position estimates, with offset correction applied.

### 6.2.4 Image Time Series

Applying the LSM technique to image sequences results in a series of image coordinates for all features and GNSS antenna centers, respectively. The time series, however, show large variabilities due to the camera movements between the individual image frames. Besides the true motion, an apparent camera movement due to the following effects is expected:

- wind (vibrations)
- changing lens arrangement due to automatic focus adjustment between the exposures
- deformation of the camera installation due to temperature variations (general temperature variations and temperature gradients due to directional heating by sunlight)

The first issue causes the external camera parameters (mainly the Euler angles) to vary. Internal camera parameters are being changed by the second effect and a mixed parameter variation is expected for the last case. All the effects are corrected by estimating the camera and additional parameters during the bundle adjustment.

As explained in Section 4.6, the GNSS antenna image coordinates are estimated by means of template matching. Fig. 6.13 shows the extracted positions of valid matches over the course of the study period. Valid matches were defined by the quality factor value ($q > 0.75$, Eqn. (3.27)), the standard deviations of the displacement parameters ($\sigma_{p_x} < 0.2$ and $\sigma_{p_y} < 0.2$, Eqn. (3.17)), as well as a rotation component of less than 10 degrees. Redundancy was increased and an estimate of the manually determined antenna center

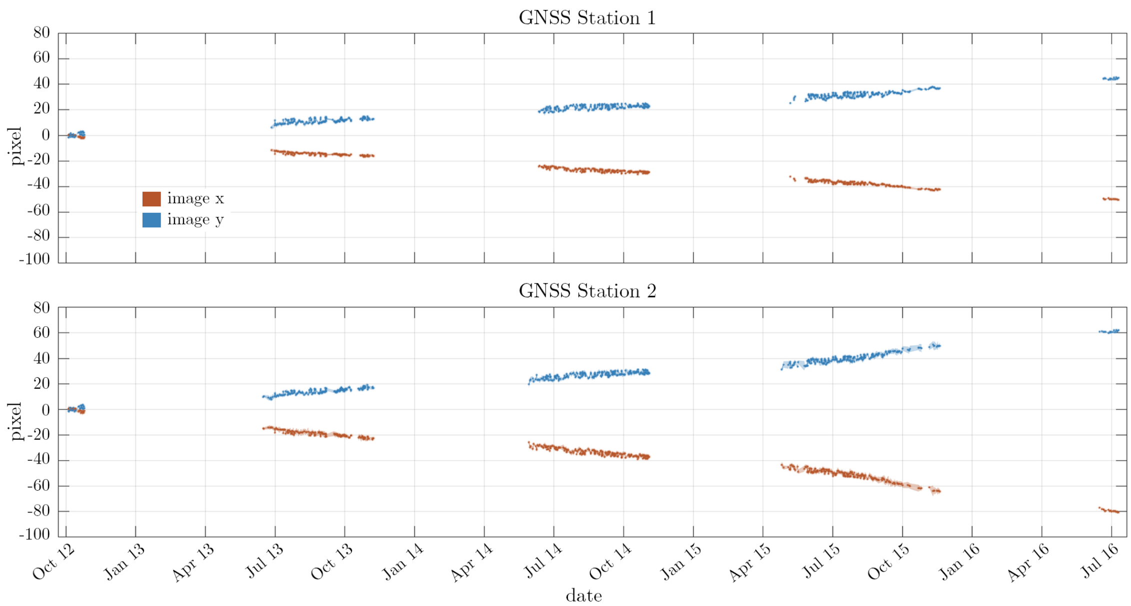


**Figure 6.13:** Relative positions of GNSS station 1 (top) and 2 (bottom) in the images seen from camera 2. The image x-coordinate is shown in brown, whereas the image y-coordinate is given in blue.

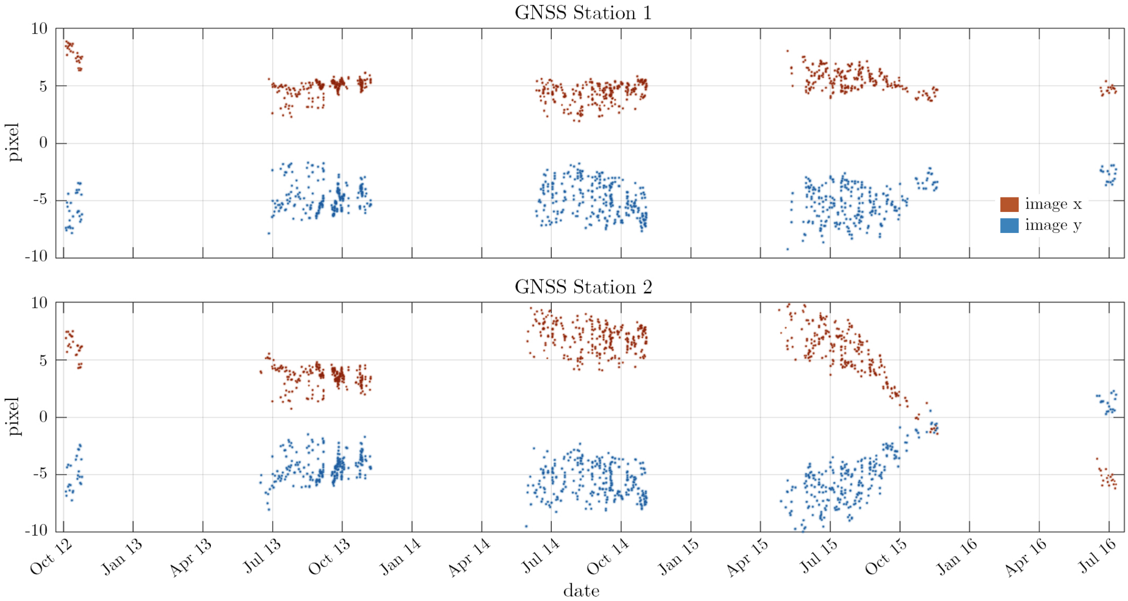


**Figure 6.14:** De-trended image coordinates of GNSS station 1 (top) and 2 (bottom) of Fig. 6.13. For the trend model, a linear regression was used. The resulting position 'noise' is mainly due to temporary camera motion (some real notable motion is observed in the summer months of 2015). Note the difference in the scale of the y-axis between this figure and Fig. 6.13.

determined by matching several templates (one for each year) with the whole image sequence. The estimated image coordinates for the antenna centers agreed within about 0.3 pixel in average.

By removing a linear trend of the station coordinates in the image, the apparent camera motion becomes evident: a random motion in the range of 2 to 8 pixel is visible (Fig. 6.14). As the true position variation is known from the 3D GNSS time series (Appendix E), a projection of these coordinates into the image space would show that the order of motion seen in Fig. 6.14 is related to the apparent camera motion.

### 6.2.5 Bundle Adjustment

Following the camera parameter determinability analysis (Section 6.2.2), all additional parameters (AP) were also included as pseudo-observations (thus acting as additional observations in the estimation procedure), whereas the high order terms ($k_2$, $p_1$, and $p_2$) were given strong weights and $k_1$ a weak weight. Parameters $k_3$, $s_c$, and $s_h$ were constrained to the initial values determined by the camera calibration. For the intrinsic and extrinsic parameters, pseudo-observations for the principal point offset ($x_p$,$y_p$), the camera constant $c$, and the camera position parameters were used. Eulerian angles $\omega$, $\phi$, and $\kappa$ were estimated as free parameters.

Whenever the following conditions were fulfilled, object point reconstruction was carried out:

- A minimum of seven ground control points are available. This guarantees that the geodetic datum is well defined and increases the probability of a good reconstruction.
- Corresponding image coordinates of points (ground control points as well as moving features) must have been successfully matched in both views.
- The camera position has to be known from the GNSS solution.

As noted in Section 6.2, SURF features are determined between the views such that for each bundle adjustment, more point correspondences could be used to determine all parameters. Although more object points have to be estimated too, the redundancy increases, thus stabilizing the estimation procedure, especially for epochs, where only a small number of LSM features was found to have successful matches.

A selection of SURF feature correspondences is found by the following procedure: putative matches are validated by estimating the fundamental matrix $\mathbf{F}$ between the views ($\mathbf{x}_1 \cdot \mathbf{F} \cdot \mathbf{x}_2 = 0$, with $\mathbf{x}_1$ and $\mathbf{x}_2$ being the homogeneous image coordinates of view 1 and 2, respectively). This is accomplished using the RANSAC procedure with a maximum allowed error of 1 pixel. To yield more positive matches, image coordinates are pre-corrected for distortion using the approximate values of the initial calibration. A variance analysis between the LSM feature matches and the SURF features showed that the LSM feature correspondences are about 4 times better in terms of precision, thus the SURF features were weighted accordingly (i.e., an observation standard deviation of $\sigma_{\mathrm{SURF}} = 0.8$ pixel was assigned).

An example of image residuals after a successful bundle adjustment for one epoch is shown in Fig. 6.15. Given the image pixel size of $5.5\,\mu$m, image residuals for the interest points (blue) are in the order of 0.2 pixel in average. Residuals of object points in larger distances tend to be smaller, whereas their absolute error increases (compare Fig. 6.11 or 6.12).

In Fig. 6.16, estimated coordinates for three selected boulders (feature *ID* numbers 48, 51, and 63 - see Fig. 6.8/6.7) are shown. For feature Nr. 48, the mean distance to the

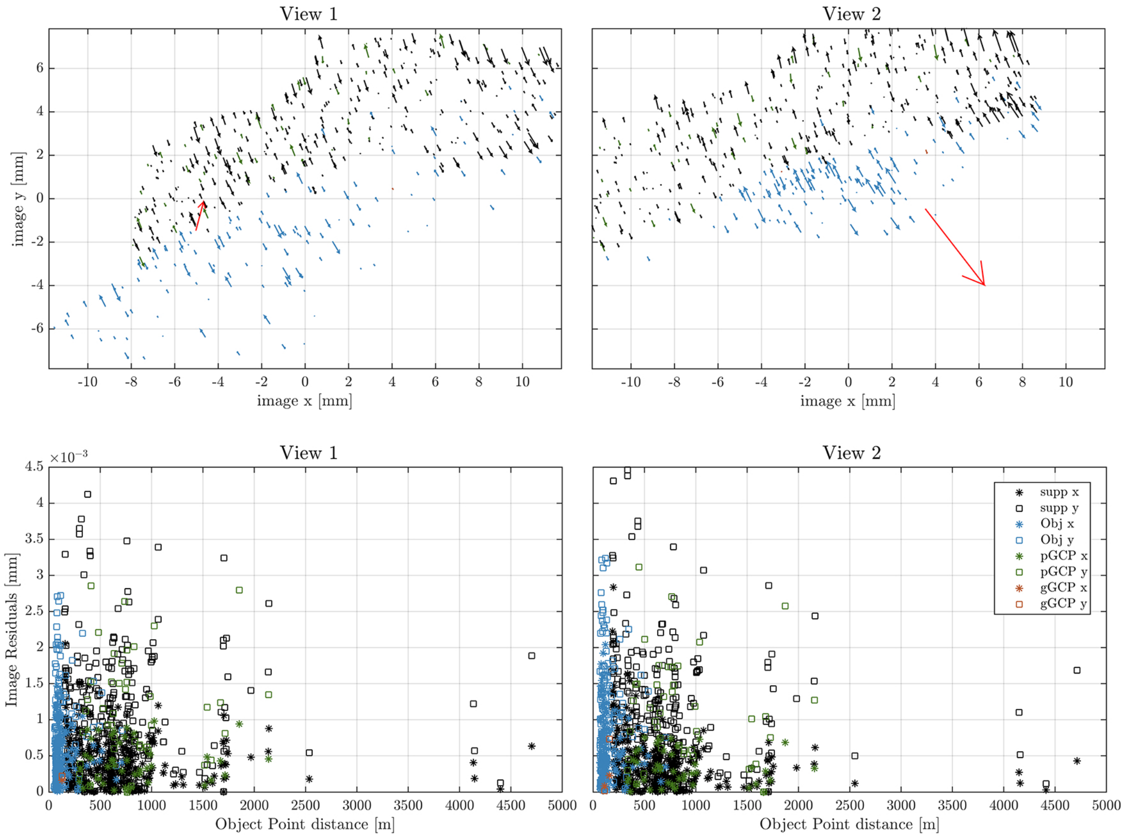


**Figure 6.15:** Image residuals in both views after a bundle adjustment. Top row of figures shows the residuals in the metrical images, amplified by a factor of 200. The bottom row of figures shows the image residuals as function of object distance. Colored in black are the supporting SURF features (different for each epoch), in blue the new object points of interest, in green the pseudo control points on solid rock, and in brown the two GNSS stations in the FOV. Gross errors are drawn in red.

two cameras measures 78 m. While this is about twice the length of the baseline between the cameras, the average $3\,\sigma$ levels are 20, 8, and 4 cm for the East-, North-, and Height components, respectively. This is also in good agreement with the varying position seen in the corresponding sequences (top of Fig. 6.16), assuming a smooth displacement as in the case of the two GNSS stations.

For the second and third bolder (*IDs* 51 and 63), respective mean line-of-sight distances of 117 and 119 m are measured. Here the $3\,\sigma$ level is generally higher than for feature Nr. 48. Although both stations are equally distant, feature Nr. 63 is located closer to the southern ridge (with respect to the camera FOVs, this is further to the right image edge). This geometrical difference results in the East component errors being almost twice as large as for position Nr. 51. North and Height components do not show a significant difference.

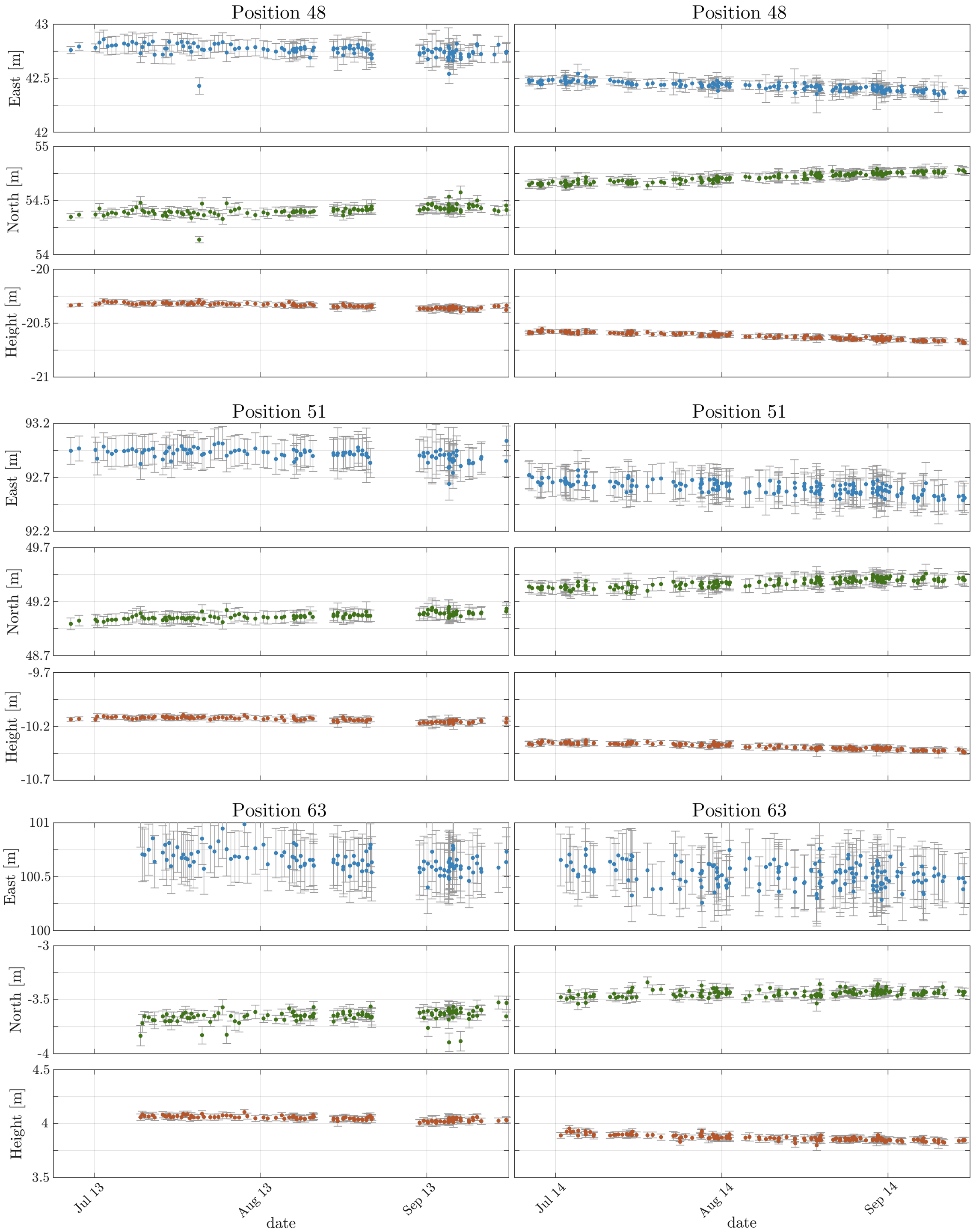


**Figure 6.16:** Example of an unfiltered position time series obtained by epoch-wise bundle adjustment. Estimates for the summer months in 2013 and 2014 are shown. Grey error bars show the $3\,\sigma$ level.

In addition to the error distribution depending on the geometry, the East component of feature 48 shows about half the error during the summer months in 2013 compared to the summer months of 2014 (Fig. 6.16). This is a direct result of the missing GNSS antenna position estimates (during the 2013 period) that was normally used as ground control point (station 2)[3]. As illustrated in Fig. 6.12, the effect of a combined solution with GNSS station 2 for the near-field features is an error reduction by approximately a factor of two. This is not the case for the far-field (i.e., $\gtrapprox 100\,\mathrm{m}$) and also not noticable in the time series for features Nr. 51 and 63 (Fig. 6.16).

### 6.2.6 GNSS for Validation

A comparison between time series of photogrammetrically derived GNSS antenna positions and the positions of the corresponding GNSS station was conducted. For this purpose, a bundle adjustment was applied without including GNSS station 1 as a ground control point (but with station 2). As demonstrated in Section 6.2.3, excluding GNSS station 1 does not significantly degrade the expected precision of the reconstruction process. Two time intervals comparing the two independent solutions for each component are shown in Fig. 6.17.

As mentioned before, no GNSS solutions for the antenna position of station 2 were available for the summer period in the year 2013. As a result, no GNSS station on the observed permafrost creep area could be used as ground control point during the summer months in this period. The estimated coordinates show a systematic error of roughly 15 cm in the East and 6 cm in the North component (no systematic offset is observed for the Height component). For the 2014 period, this offset has vanished as GNSS station 2 could be integrated in the computation, again showing the advantage of using a well located high precision reference point. Within the estimated position errors, relative displacements show no systematic drift.

[3] A second, non-concentric GNSS antenna is mounted on the same mast. Although position estimates are available for this period, those were not used in this work. In case the rock boulder does not have a rotational component in the z-axis, a simple offset correction could have been applied to fill the measurement gab of the concentric antenna. However, because such a rotation can not be estimated with a single position measurement and also because the direct identification of the second antenna in the image is more difficult, the non-centric position solution was not used.

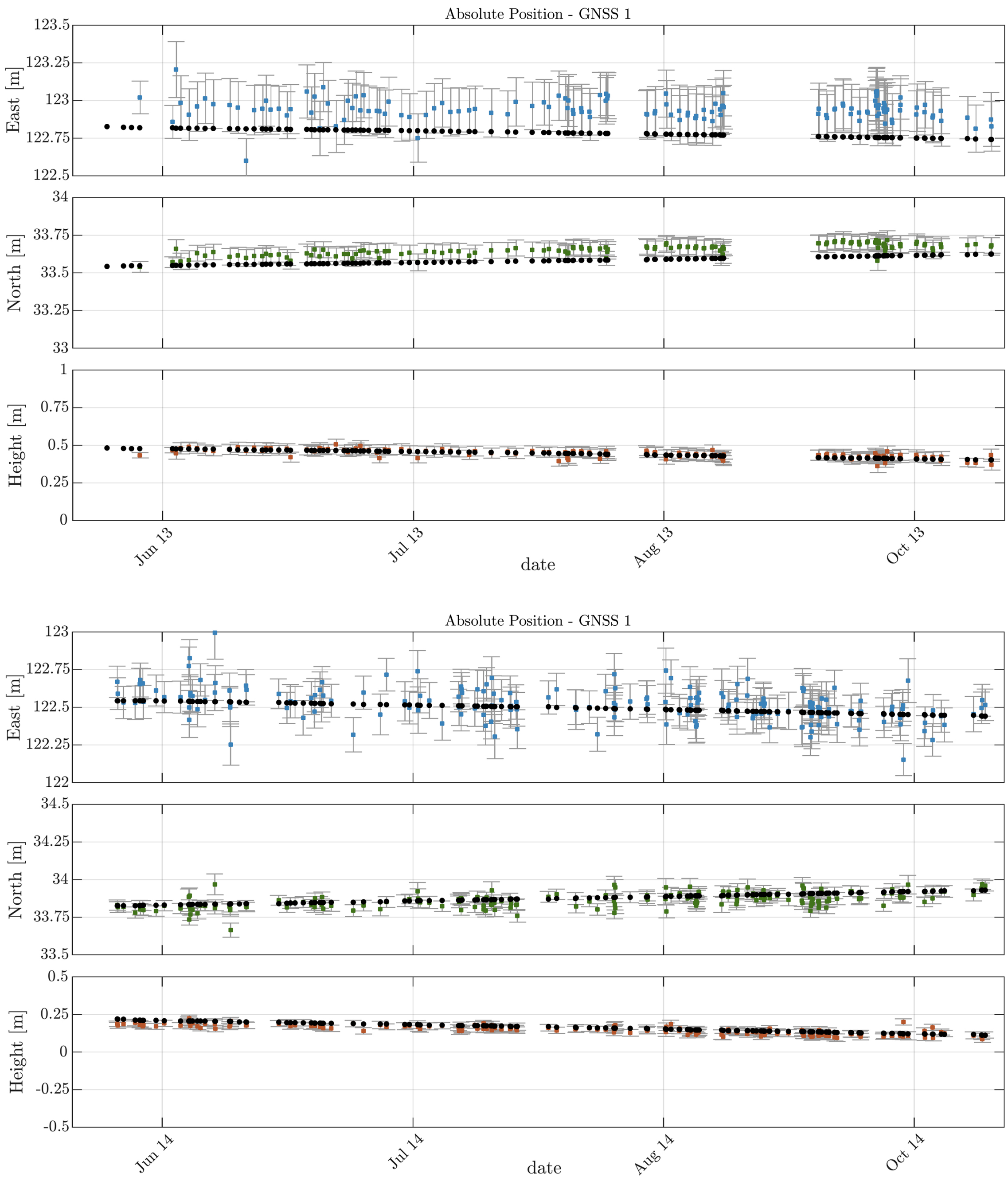


**Figure 6.17:** Difference between GNSS and photogrammetric position estimation for the summer months of 2013 (top) and 2014 (bottom). Photogrammetric coordinates are presented in color, whereas the GNSS reference is shown in black.

### 6.2.7 Annual Displacements

Given the slow movement of the observed permafrost creep, yearly displacement rates are summarized. The increased accuracy of the general yearly flow behavior also allows to recognize displacement trends more easily. The average displacement over the complete measurement period is shown in Fig. 6.18.

Regional yearly displacement rates are given in Tab. 6.3 with zones defined in Fig. 6.19. Areas were defined along the main flow direction, whereas the front was divided into another two zones, given the high density of features and based on the result presented in Section 6.4. The general displacement rate over the course of the three years 2013, 2014, and 2015, indicate an acceleration. Absolute (3D) velocities in 2015 are about 50% higher than in the year 2013. This is mainly due to the acceleration phase observed in the late summer months in 2015 (see Section 6.3), during which velocities increased by about a factor of two on average.

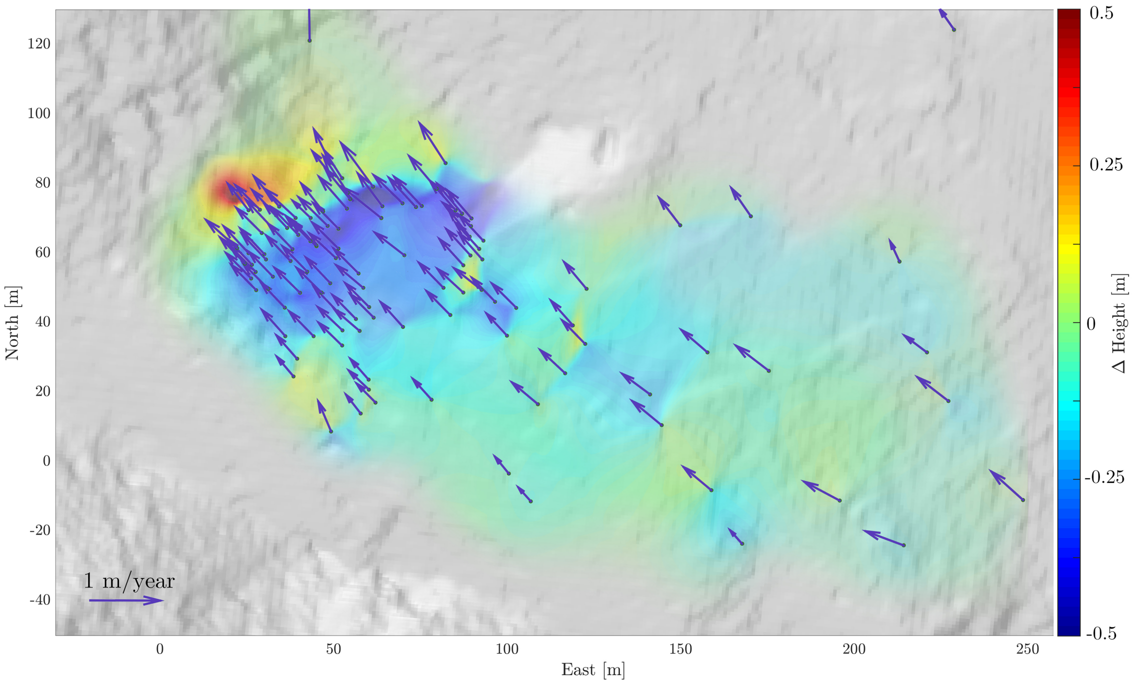


**Figure 6.18:** Mean displacements measured between October 2012 and July 2016. Height changes were interpolated on a regular grid and are colorized on the shaded DEM relief (provided by FOEN). The differences in Height are due to a (dominante) translation of the moving mass having an irregular topography.

**Table 6.3:** Average yearly displacement rates given for the zones defined in Fig. 6.19. Values were computed as the weighted mean of all displacements estimated during the respective time intervals being in the corresponding zones. Average errors are in the order of 0.07 m/year for the East and 0.03 m/year for the North and Height components, respectively, whereas estimated errors for the two GNSS stations G1 and G2 are smaller by about two orders of magnitude. The average yearly displacement rates are given in m/year and were computed using the collocated position differences (see Section 6.3) between January and December of the respective year.

| Zones & GNSS | component | 2013 | 2014 | 2015 |
|---|---|---|---|---|
| Z1 | East | −0.405 | −0.409 | −0.409 |
| | North | 0.194 | 0.248 | 0.314 |
| | Height | −0.263 | −0.299 | −0.405 |
| | 3D | 0.520 | 0.564 | 0.656 |
| Z2 | East | −0.314 | −0.350 | −0.372 |
| | North | 0.296 | 0.318 | 0.383 |
| | Height | −0.245 | −0.321 | −0.376 |
| | 3D | 0.496 | 0.572 | 0.653 |
| Z3 | East | −0.310 | −0.332 | −0.409 |
| | North | 0.281 | 0.361 | 0.365 |
| | Height | −0.190 | −0.296 | −0.340 |
| | 3D | 0.460 | 0.573 | 0.645 |
| Z4 | East | −0.270 | −0.423 | −0.471 |
| | North | 0.361 | 0.405 | 0.621 |
| | Height | −0.259 | −0.325 | −0.456 |
| | 3D | 0.512 | 0.670 | 0.903 |
| Z5 | East | −0.303 | −0.358 | −0.445 |
| | North | 0.288 | 0.369 | 0.445 |
| | Height | −0.234 | −0.325 | −0.383 |
| | 3D | 0.479 | 0.608 | 0.737 |
| GNSS 1 | East | −0.248 | −0.318 | −0.405 |
| | North | 0.259 | 0.318 | 0.405 |
| | Height | −0.234 | −0.310 | −0.383 |
| | 3D | 0.428 | 0.546 | 0.689 |
| GNSS 2 | East | −0.281 | −0.409 | −0.551 |
| | North | 0.358 | 0.507 | 0.788 |
| | Height | −0.296 | −0.427 | −0.617 |
| | 3D | 0.543 | 0.779 | 1.143 |

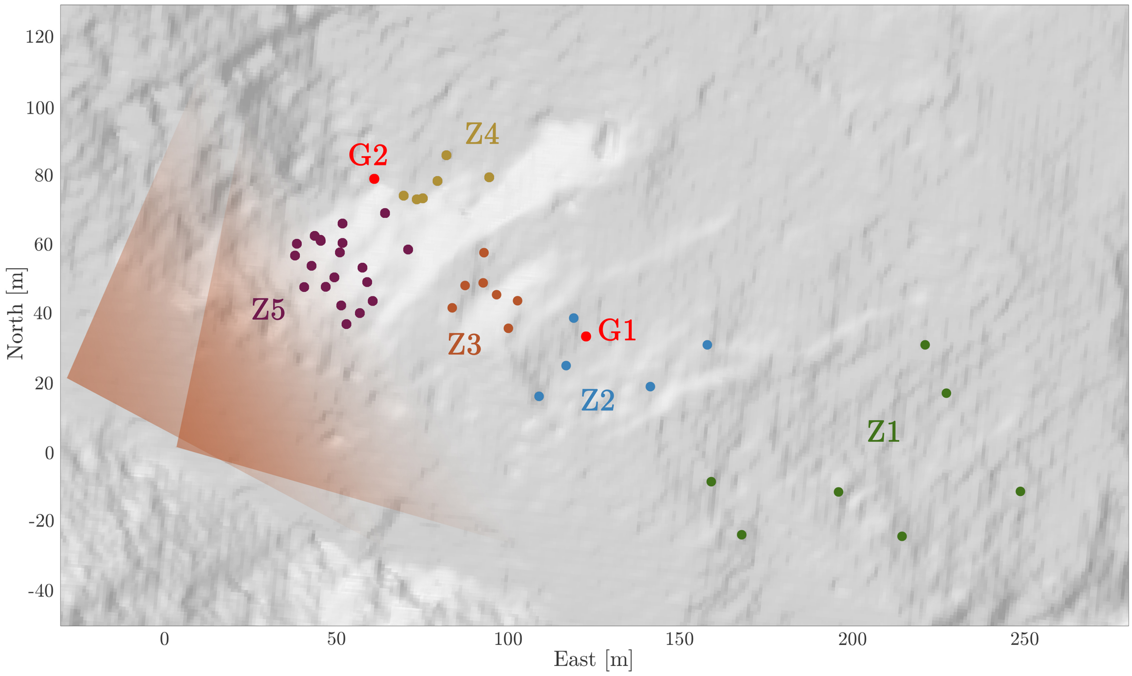


**Figure 6.19:** Definition of zones $Z1$ to $Z5$ as used in Tab. 6.3. Brown shadings indicate the respective camera FOV.

## 6.3 Collocation

Typically, the principle quantity of interest for surface displacement monitoring applications is velocity as a function of time and space, $\mathbf{v}(\mathbf{r}, t)$. This quantity can be derived by the collocation technique principle, introduced in Section 5.1, using the estimated feature position coordinates for the various epochs as observations.

Correlation functions, for each feature, were empirically determined using the technique described in Section 5.2 with a relative bin size of 0.5 [-]. The estimation principle followed a robust least-squares adjustment (compare Section 3.2.6) that turned out to yield more realistic results in some cases than the ordinary least-squares approach. For every feature, parameters of model 1 and 2 (Eqn. (5.8) and (5.9)) were estimated. The model that better matched the data (lower a posteriori variance) was chosen[4]. Four examples of empiric covariance functions for different rock boulders are shown in Fig. 6.20. On average, the exponent $n$ (Eqns. (5.8) – (5.8)) was found to be 2 for almost all cases, whereas the correlation lengths, $r_s$, were found to be between 30 to 120 days. This large scattering can be explained by the differences in the deterministic models used (see below) and eventually also the differences of the natural motion and noise content in the data.

[4]model 3 was tested as well but due to the nature of the harmonic function, unrealistic oscillations were observed in periods of poor data coverage. Therefore only model 1 and 2 were considered

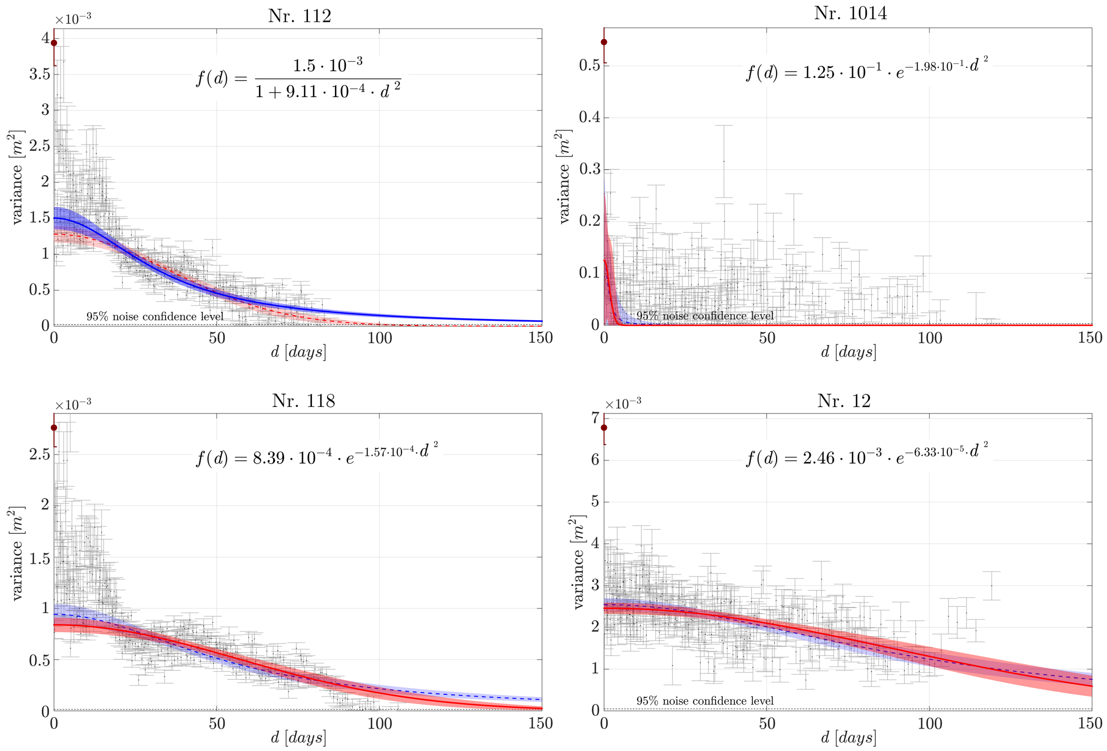


**Figure 6.20:** Examples of empirically determined covariance functions. The numbers correspond to *IDs* defined in Figs. 6.8 and 6.7. Nr. 1014 is a pseudo ground control point with no motion and no significant temporal correlation, as seen in the figure. Grey error bars show the errors of the individual covariogram components. Covariance function defined as model 1 (Eqn. (5.8)) is shown in red, where the blue curve shows the covariance function following model 2 (Eqn. (5.9)). A deterministic model of degree 1 was used in all cases. Total variance is indicated by the brown mark at $d = 0$.

Collocation was applied for each time series, starting in October 2012 and ending in July 2016. A solution was computed (predicted) for every day within this period (1378 days in total). Applying the principles of collocation to irregularly sampled data, however, can be critical, especially if the covariance function shows a correlation length that is in the same order as the continuous sequences in time (as it is the case here). The choice of the deterministic model then plays an important role as it also effects the correlation length and the ratio between the stochastic signal and the noise components. Thus a suitable higher order deterministic model is preferred as correlation length thereafter generally decreases. For the current scenarios, polynomials of degree 2 or lower were used.

The sensitivity of the chosen deterministic model, i.e., polynomials of varying degree, w.r.t. the empirically determined covariance function is illustrated in Fig. 6.21. Polynomials of degree 0 (case (1)), 1 (case (2)), and 2 (case (3)), for a position time sequence were used for de-trending (using trend functions $f(d)_{\mathrm{det}}$), before the parameters of the covariance functions were determined. The respective correlation lengths found for hereby are 162, 33, and 38 days. For case (1), a high variance ration $r = \sigma_s^2/\sigma_n^2 = 3.6$ indicates a high signal

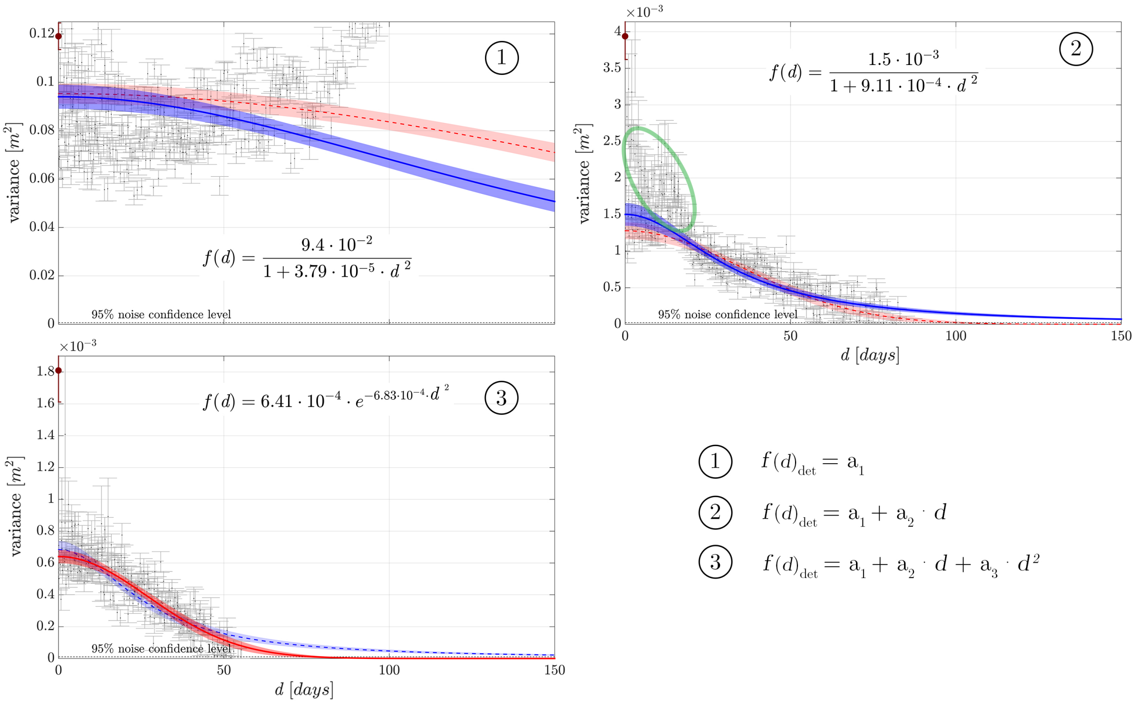


**Figure 6.21:** Difference in the estimated covariance functions w.r.t. the chosen deterministic model. The time series of *ID* Nr. 112 was used here. Coefficients $a_1$, $a_2$, and $a_3$ are the parameters of the deterministic model. The brown mark at $d = 0$ indicates the total signal variance, i.e., $\sigma_0^2 = \sigma_s^2 + \sigma_n^2$. With increasing model complexity, differences between the measurements and the detrended measurements become smaller and so does the total signal variance. The green ellipse shows an obvious disagreement between the functional fit and the data points (grey) in the near field ($\leq$ 20 days).

content in the de-trended measurements and the functional fit (here the blue curve) is a rather poor representative of the date (grey), i.e., a large scattering is seen. A better fit is obtained for case (2), with a polynomial of degree 1: the differences between the data points and the analytical function are considerably smaller than before, specially for time differences larger than 20 days. A relatively large scattering remained for the close range ($\leq$ 20 days, see green ellipse in Fig. 6.21) that was, however, partially ignored during the robust least-squares adjustment of the covariance function. Increasing the polynomial degree to 2 (case (3)), the evident near-field scattering vanished. The variance ratio $r$ is about 0.6 for case (2) and (3), and also the correlation lengths are in good agreement between these two models. The total signal variance in (3) is by a factor of two lower than for case (2), indicating a better overall fit of the trend model. Note that the vertical scale between the three graphs is different and therefore differences between the estimated covariance function and the data points can not directly be compared. Using polynomials with higher degrees did not show any improvement but unfavorable effects of an over-parameterization.

Practically, the differences between collocated position coordinates and velocities are small, when comparing results obtained by using a polynomial of degree 1 and 2 as the determin-

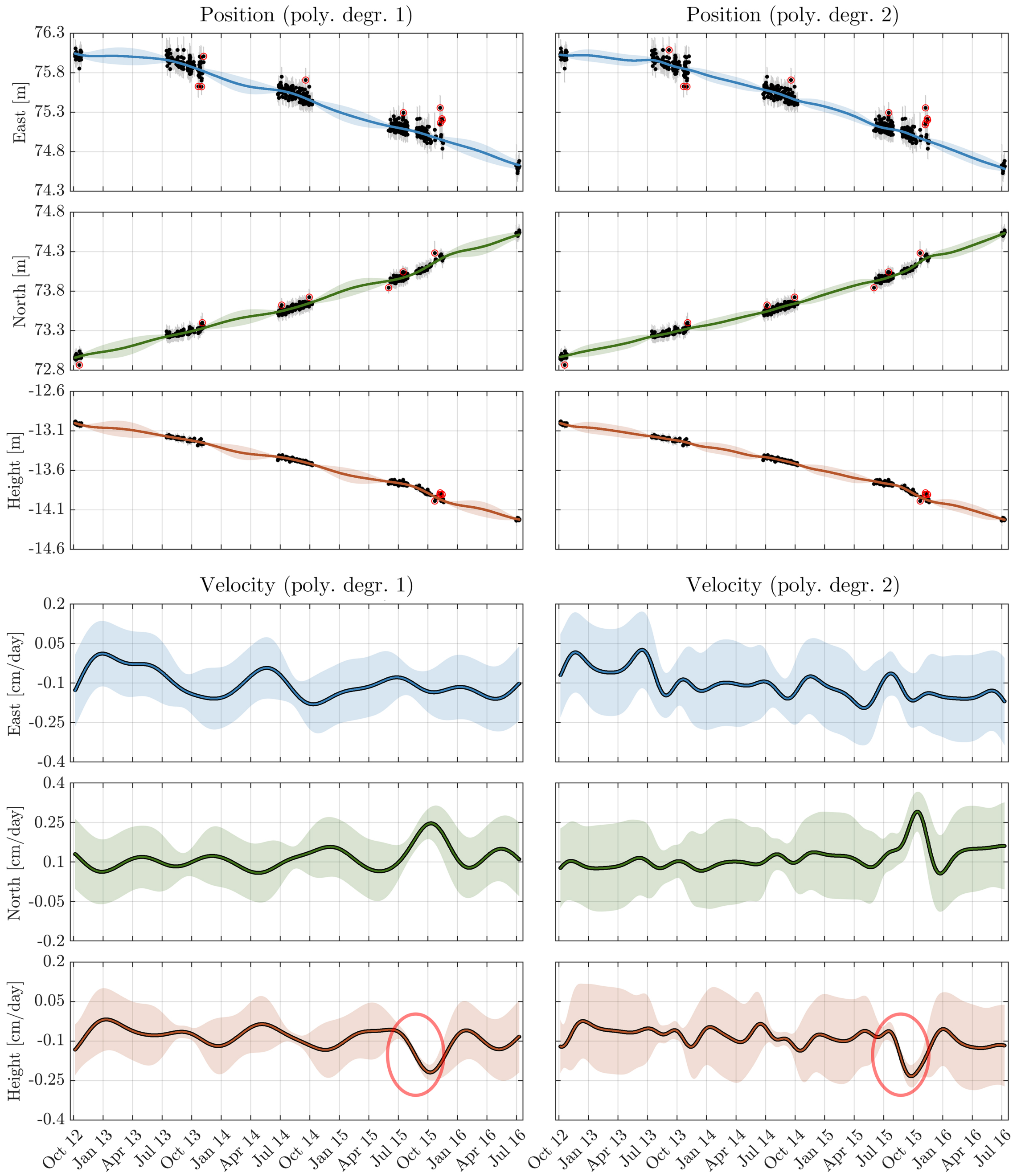


**Figure 6.22:** Time series of feature Nr. 78. The top row shows the collocated position coordinates (with unfiltered data as black dots and outliers in red) while the respective velocity estimates (computed using the principles described in Section 5.1.2) are the shown in the bottom row. The left column shows the results obtained when using a polynomial of degree 1 for the deterministic model and the right column illustrates the respective position and velocity estimates when using a polynomial of degree 2. A noticable difference is highlighted by the red ellipses.

istic model. Fig. 6.22 shows this effect for feature *ID* 78. In this example, the correlation lengths for the degree 1 and degree 2 polynomials were 101 and 55 days, respectively. While there are hardly any differences in the estimated position solutions, the velocity shows a faster acceleration where data coverage is good (see red ellipses in Fig. 6.22) and larger small scale disturbances (along with larger formal errors) where data coverage is poor. As seen for the velocity estimates in the North component, the peak around October 2015 is found to be larger when using the higher order polynomial. These effects are directly related to the difference in the estimated correlation lengths. For larger correlation lengths, the collocated position and velocity estimates are farther passed into periods of missing data, comparable to a larger smoothing window. In contrast, smaller correlation lengths cause a higher sensitivity to temporal differences.

With the result presented above, a polynomial with a degree 2 was preferred as the trend function. In some cases of poor data coverage, degree 1 was used instead. Reliable collocated results are obtained only in areas of good data coverage, as shown in the examples above. Accuracy estimates for the predicted position coordinates (and velocity) during the winter months are dominated by the deterministic trend error and might thus be heavily underestimated.

The collocation principle was applied to the time series obtained by bundle adjustment as well as to time series of feature displacements in the image space (Section 6.1). For each coordinate sequence, (3D) and (2D), the following procedure was carried out:

(a) Definition of the deterministic model based on the distribution and number of available measurements. The principle here is to assign a simple model in case only very few measurements are available or if they are not regularly distributed along the time interpolation period.

(b) Reduction of position coordinates by removing the deterministic model part.

(c) Estimation of the covariance function with its associated parameters using the reduced position coordinates. A single covariance function is estimated using all dimensions of the measurements (i.e., East-, North-, and Height-components for the 3D trajectories or image x and y directions for the single view analysis).

(d) Applying the principles of collocation.

(e) Detection of outliers in the residual noise component by a thresholding approach. A large variance is assigned to the detected erroneous observations.

(f) If erroneous observations have been detected, the sequence (d)-(e) is repeated until no more outliers are detected.

(g) Computation of variance-covariance components of the interpolated signal and of the derived velocity components.

A typical example (i.e., feature *ID* 99 at an average distance of 70 m from the cameras) of a collocated position time series for the acceleration period in September 2015 is given in Fig. 6.23, and for the full time period in Fig. 6.24. Error propagation for the collocated position estimates predicts an average error of 3 cm for the East, 1.5 cm for the North, and 0.8 cm for the Height component, respectively, whenever measurements are present. These error propagation results, however, strongly depend on the estimated covariance function, as described before, and therefore need to be considered with care. The estimated velocity components, with corresponding $3\,\sigma$ envelopes, show an average velocity of -0.1, 0.1, and -0.1 cm/day for the East, North, and Height components, respectively. As seen in Fig. 6.24, error propagation for the three velocity components results in $3\,\sigma$ levels of approximately 0.10, 0.06, and 0.04 cm/day for the measurement periods (summer months), and more than 0.10 cm/day for prediction periods (winter months) for all components. Again, these estimates heavily depend on the chosen covariance function and on the relative location between the observed feature and the two cameras. In addition to the general fluctuation of the velocity estimates, a long term acceleration can be observed over the complete measurement period, mostly in the East and Height component.

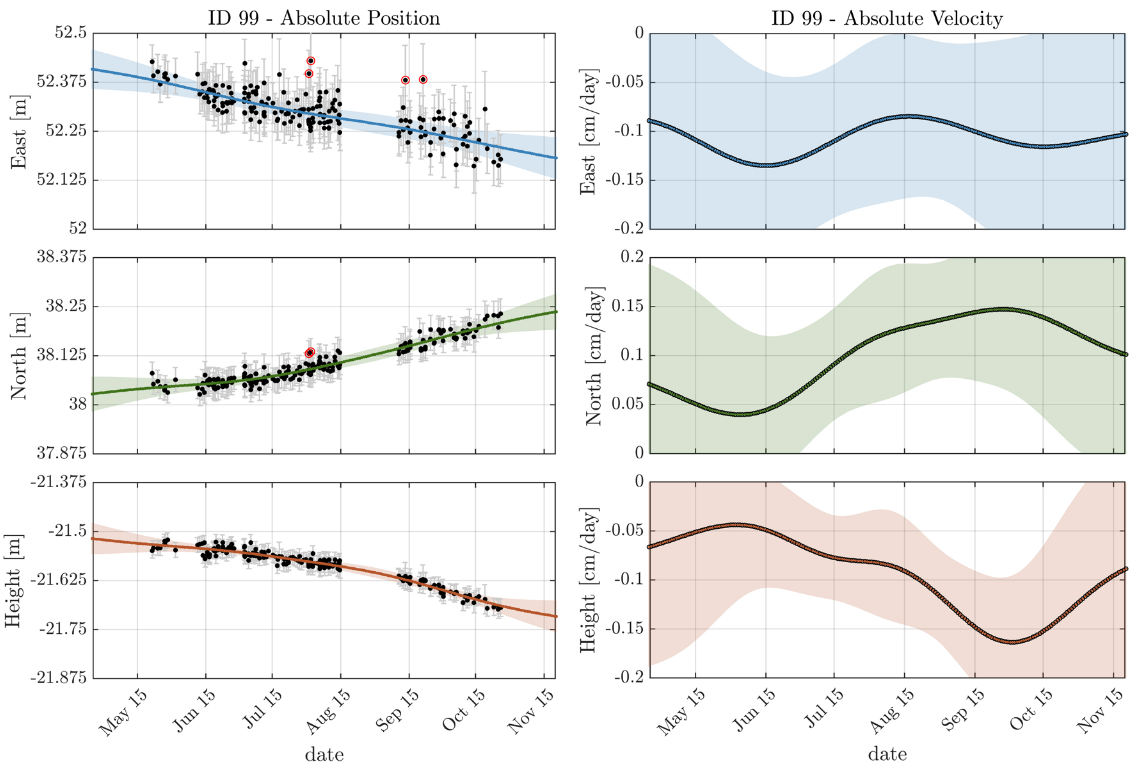


**Figure 6.23:** Acceleration phase that started around August 2015. The vertical axis for the position coordinate estimates spans 50 cm in all components. Shaded envolopes show the $3\,\sigma$ level. The full sequence of the time series is given in Fig. 6.24.

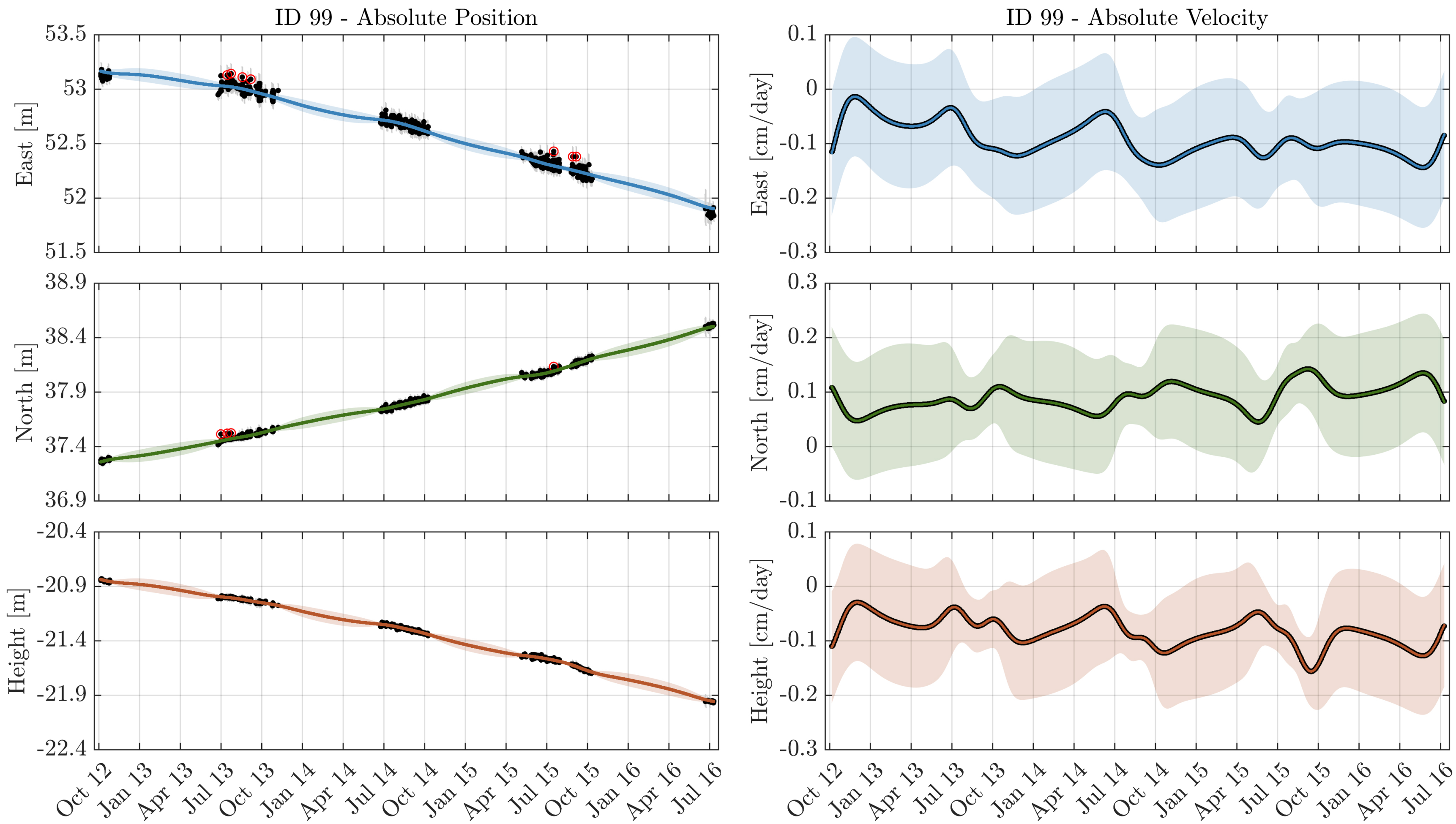


**Figure 6.24:** Unfiltered and collocated position coordinates from feature Nr. 99 on the left and derived velocities on the right. Error bars show the $3\,\sigma$ level. Position estimates marked in red were excluded in the collocation procedure (outliers).

## 6.4 Adaptive Collocation

This section presents results obtained by applying the principles of adaptive least-squares collocation (Section 5.3). 2-dimensional flow fields obtained by bundle adjustment (East and North component) were used as primary observations. Velocities were estimated for each time sequence using the principles described in Section 5.1.2 with results presented in Section 6.3.

The following parameters were used for the initial step, $k = 0$:

- correlation length of the inhomogeneous signal $r_{s_i}^{k=0} = 40\,\mathrm{m}$
- correlation length of the homogeneous signal $r_{s_h} = 15\,\mathrm{m}$
- variance of the inhomogeneous signal $\sigma_{s_i}^2 = \sigma_0^2 - \sigma_n^2$
- variance of the homogeneous signal $\sigma_{s_h}^2 = 0$
- dilation length $\gamma = 3\,\mathrm{m}$
- dilation scaling $\lambda_{max} = 1.25$

with $\sigma_0^2$ being the variance of the signal after reducing the measurements by the deterministic trend, and $\sigma_n^2 = \sigma_v^2$, i.e., the variance of the estimated velocity components. For the deterministic model, a polynomial of degree 0 for both components has been chosen because of the nearly homogeneous flow field (i.e., mean velocities $v_x$ and $v_y$ are determined).

After the first iteration, $\sigma_{s_h}^2$ is adjusted such that the norm of the residual noise components is minimal. During this (separate) iterative correction of $\sigma_{s_h}^2$, $\sigma_{s_i}^2$ is adjusted such that $\sigma_{s_i}^2 = \sigma_0^2 - \sigma_{s_h}^2 - \sigma_n^2$. Because the total dilation defining the new metric is a sum of dilations applied in each grid point (Eqn. (5.23)), the dimension of the grid increases in general. To compensate this average increase, the correlation length ($r_{s_i}^k$, Eqn. (5.16)) is adjusted, i.e., increased, after iteration. The correction factor is computed by the ratio of the mean *before* and *after* neighborhood distances at the grid position of maximum dilation. Typical correction factors obtained hereby are in the order of 1.1 for each iteration.

Fig. 6.25 shows the changing (grid) metric for each iteration (eight iterations were applied in total). Starting from a regularly spaced grid (~3 m spacing), a prominent dilation occurs mainly in the North-South direction. The collocated velocity field is shown in Fig. 6.26: given are the collocated displacement rates for the year 2014 after the first iteration, i.e., the traditional least-squares collocation solution (a), the result obtained for the adaptive collocation in (b), and the difference between the latter two (amplified by a factor of 10). The latter plot shows that the ALSC result - although very similar to the LSC in this case - estimated velocities in the front region being about 0.03 m/year faster as compared to the LSC result. Other prominent differences are found at the southern edge of the motion field (along the Northwest orientated boundary).

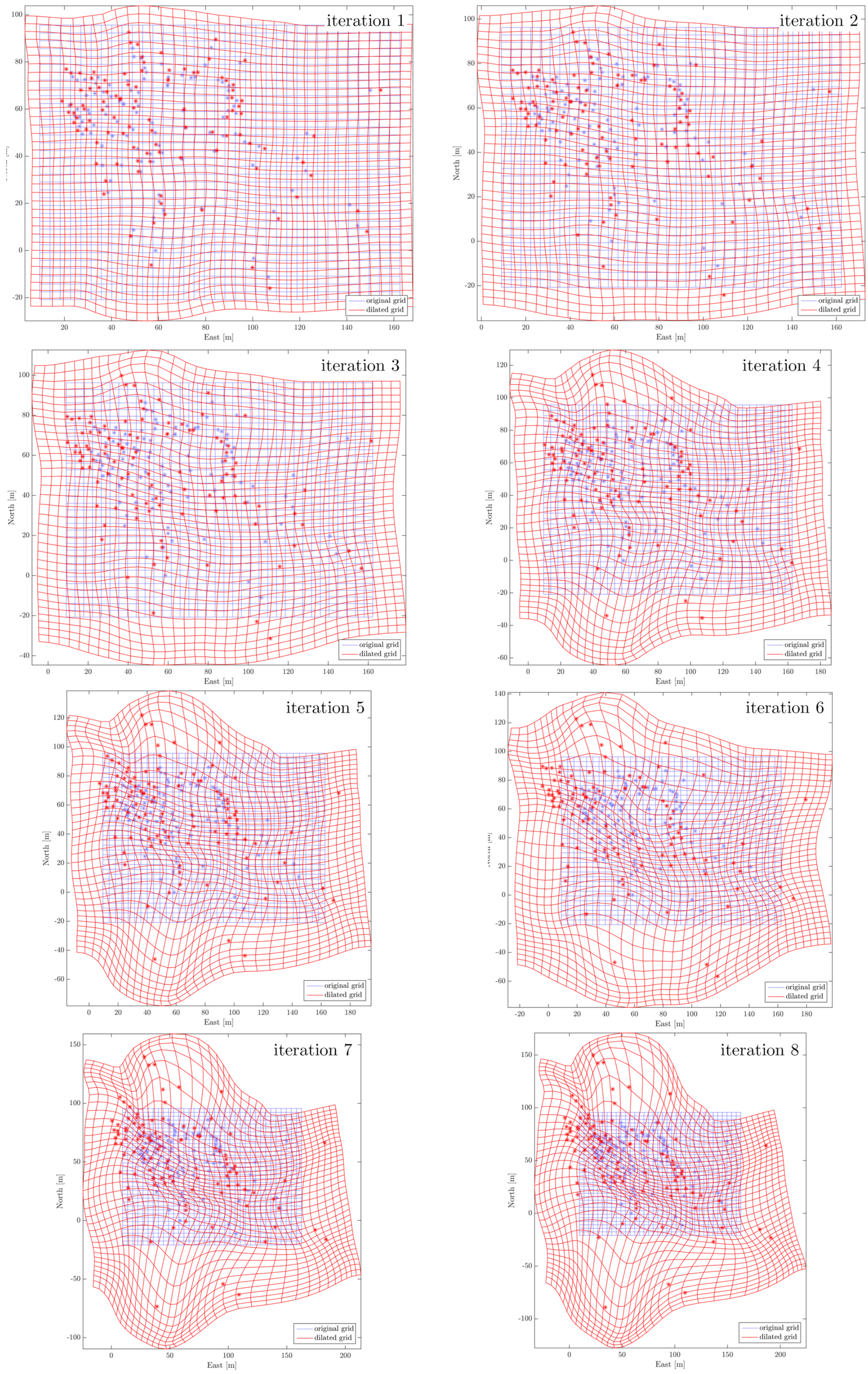


**Figure 6.25:** Stepwise dilation of the metric grid during the ALSC procedure for the estimated displacements in the year 2014. A total of eight iterations were applied.

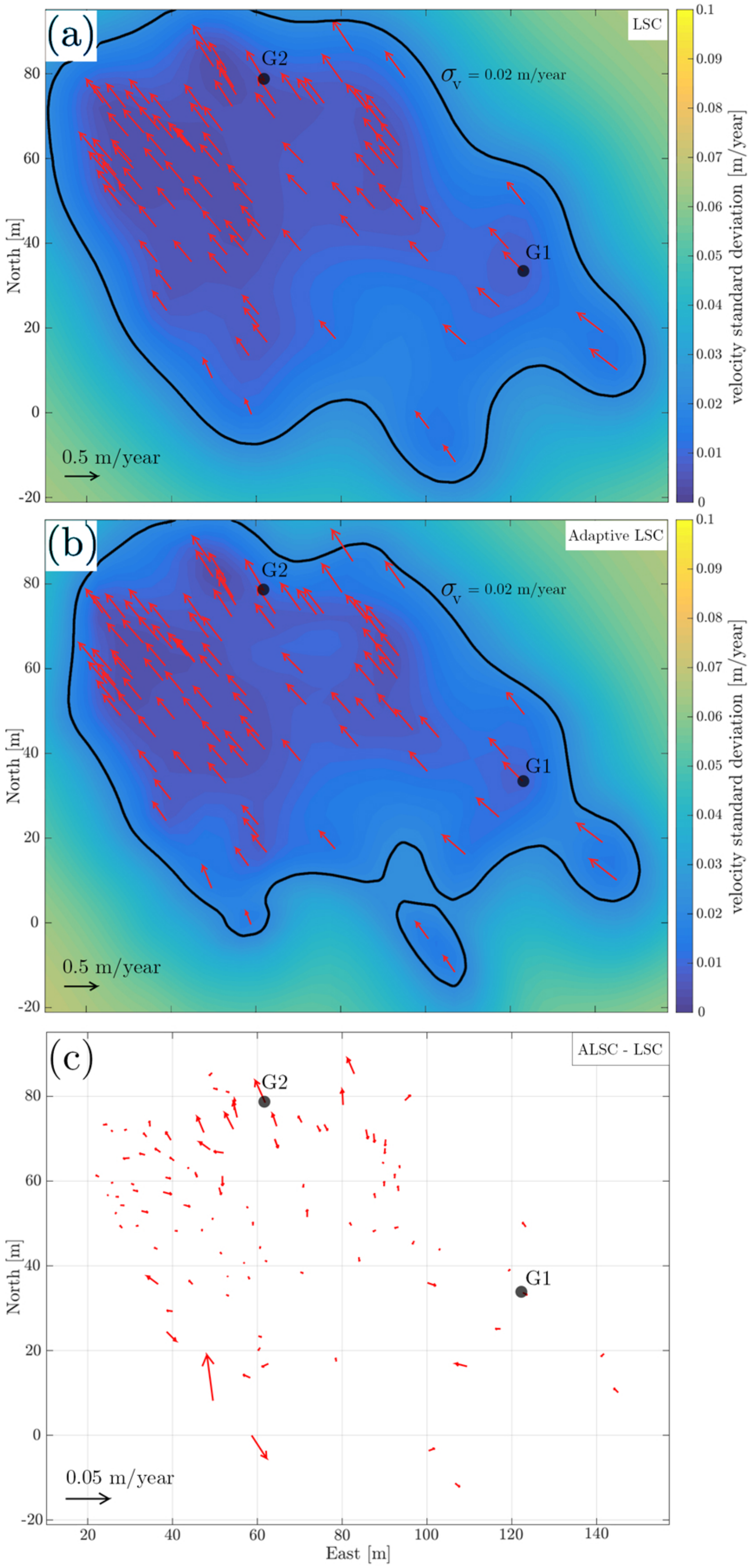


**Figure 6.26:** Collocated velocity fields for the year 2015. (a) shows the least-squares collocation result using a single homogeneous covariance function (i.e., first iteration of the ALSC process). (b) the adaptive least-squares result after eight iterations. (c) the difference between (a) and (b), amplified by a factor of 10. Black markers indicate GNSS stations G1 and G2 and the black contour encloses areas with accuracies better than $\sigma_v = 0.02\,{}^{\mathrm{m}}/_{\mathrm{year}}$.

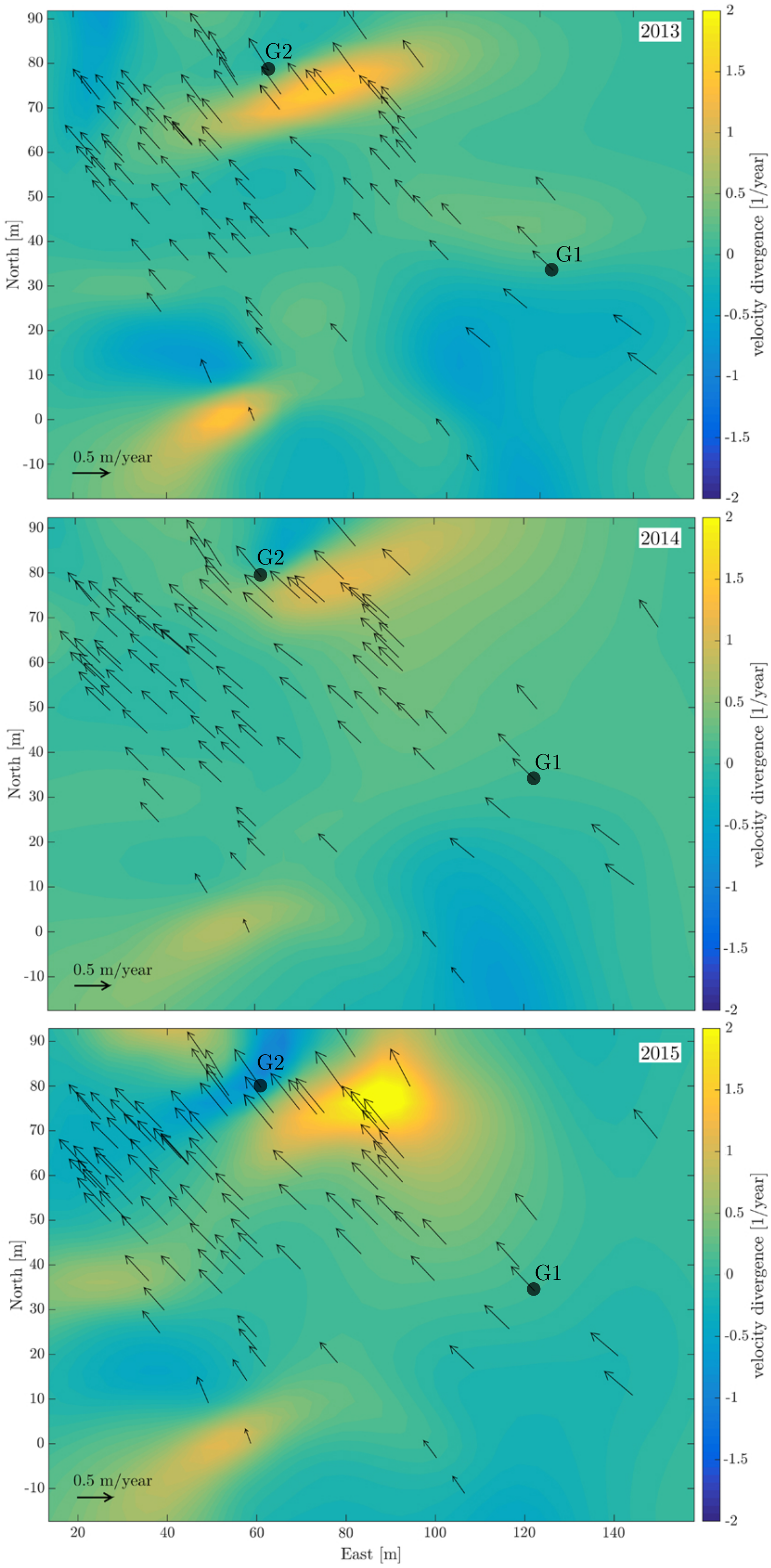


**Figure 6.27:** Divergence of the velocity field (components East and North), iteratively determined during the ALSC. Black markers indicate GNSS stations G1 and G2. Areas in yellow show zones of extension while areas colored in blue indicate compression areas. The location of the prominent extension area is at the edge of the observed rock boulders so that its exact location is not very well constraint and varies slightly over the years. Arrows show the mean annual displacements.

As described in Section 5.3.3, the dilation process is estimated using the velocity gradient field as the driving force. Applying the correction factors (see Section 5.3.4) to estimate the strain rate w.r.t to the original (orthogonal) metric, the velocity divergence field (computed as the trace of the strain rate tensor for each prediction point) can be used to detect areas of extension and dilation. Fig. 6.27 summarizes the results found hereby: two areas are affected by a dilation process (positive values, i.e., yellow), also remaining stationary over the course of the three years 2013, 2014, and 2015. The first area, with a center at approximately 80 m East and 75 m North (in the local topocentric coordinate frame), shows an extension (decoupling) with a significant increase in 2015. The second area in the lower half of the figures shows another extension region at the edge of the measured velocities.

It is to note that the ALSC procedure applied to this series of 2-dimensional flow fields is sensitive to the initially chosen ALSC parameters. As this only shows the horizontal flow components, a full picture of the scene can only be obtained by including the full 3-dimensional measurements. For such a scenario, however, a suitable model to handle the free surface in the 3D volume as well as assumptions or physical conditions for penetration in depth are needed.

# 7 Conclusions

This thesis combines the image-based photogrammetric processing techniques with high precision GNSS positioning solutions for an accurate estimation of motion at various positions on the Grabengufer rock slide, Mattervalley VS, Switzerland. A stereo-pair of prototype online camera systems, equipped with commercial digital single-lens reflex (DSLR) cameras, embedded PC platforms, and GNSS sensors, were constructed and deployed in the vicinity of the permafrost area. This work was realized by the project partner (Computer Engineering Group, ETH Zürich) that also developed low-cost GNSS stations (using L1-GPS frequency receivers), of which two were installed on moving rock boulders within the field of view (FOV) of the cameras.

The thesis comprises three main topics, the first focusing on an accurate feature tracking procedure that is used to estimate image coordinates of objects between various image pairs. In the second topic, these image coordinates are used to estimate displacements, either by combining a camera with a digital elevation model (DEM) or by scene reconstruction using the stereo-pair cameras. The principles of collocation were explored in the third topic, aiming at an adequate filtering and prediction procedure. Given the outline of these topics, the main results and critical aspects are discussed below.

### Feature Tracking

Over the course of this work, a semi-automatic processing pipeline was developed to select images suitable for motion analysis. Short summer periods at these high altitudes (2'900 a.m.s.l.), and therefore long periods of snow coverage, limit the time window for the permanent observation of rock movements by optical imagery. Given these conditions, the time span for selecting suitable images was chose such that periods of partial snow coverage were also included. An elementary classification approach was implemented for the detection of areas covered by snow that were treated as unusable areas in the image matching process. Without this pre-analysis, the systematic retreat of snow during the melting period would have caused a degradation of the matching results.

The method of least-squares image matching (LSM) was chosen as feature tracking method. A flexible implementation of this matching technique was conducted and used for different tasks: (a) for the automatic tracking of natural features between various image pairs, (b) for matching GNSS image templates to recover the image coordinates of the GNSS antennas, and (c) for matching natural features between the views of the cameras. Testing the accuracy of the method, empirical standard deviations clearly below 0.05 pixel for the translation components were found in case the images were contaminated by synthetic random noise. Robustness to gross errors was examined in a second test, showing an empirical precision of $\pm 0.05$ pixel in scenarios with up to 40% gross error coverage. For matching features exposed to various lighting conditions, combined with model errors and other uncertainties, a matching accuracy in the order of 0.1 to 0.5 pixel can be expected - also confirmed in the reconstruction process.

A strategy was proposed to iteratively improve the matching result by a dynamic adjustment of the matching window size. The automatic procedure uses the structural content of the image near the feature of interest to evaluate an optimal matching window while giving also the possibility of a pre-defined minimal dimension. The latter is useful to prevent inaccuracies (model errors) due to the various local lighting conditions. The proposed matching strategy was found to have a success rate of more than 90% for features seen over a period of two years.

### Reconstruction

The estimation of permafrost displacement rates was conducted, in a first step, by the principles of monoplotting: LSM was applied to the sequence of images seen by a single camera. The resulting chronologies of image coordinates, each representing an specific feature, were corrected by an image rectification process needed to compensate the apparent camera motion: wind, temperature variation, and focus adjustments caused the image FOV to vary randomly by up to 8 pixel. Necessary correction parameters are typically estimated by minimizing position differences in stable areas around the moving object. Here, however, an unfavorable geometry of such stable areas seen by the camera did not allow for a conventional accurate estimation of these parameters. To carry out the monoplotting procedure, image coordinates were rectified using the camera parameters determined during bundle adjustment. Error propagation was conducted for the sequence of parameter estimations involved: initial image coordinate accuracies estimated by the LSM procedure were combined with the covariance matrices of camera parameters for the rectification process. The level of standard deviation in the resulting image coordinates was found to be in the order of 0.5 to 1.5 pixel. These error levels heavily depend on the geometrical setting and are different for other scenarios. For the present case, i.e., the stereo-view geometry, the quality of the cross-view feature matches and the precision of ground or pseudo ground control points was found to have a major influence on the estimated camera parameters because the parameter variance-covariance matrix is a result of the joint system.

Adjusted image coordinates were further used for the collocation process, yielding collocated position and velocity time series. The last step, i.e., scaling the velocity estimates to metrical units, was performed by the principles of ray tracing: a scaling factor was determined based on the distance between the camera origin and the intersection point of a ray hitting the surface of a recent DEM (provided by the Federal Office of the Environment). The direction of each ray is expressed as a function of the camera origin and orientation as well as the position of the object in the image. Because formal error propagation results can be misleading for the present geometry, a Monte Carlo simulation was performed to obtain more realistic error estimates. While differences are usually not significant for aerial photogrammetric tasks (camera is typically nadir viewing), it was shown that the present oblique geometry has a strong impact on the validity of the scaled displacement estimates that is due to a heavy underestimation of the formal uncertainties in some image areas.

A direct 3-dimensional reconstruction was performed for the various epochs by the principles of bundle adjustment. Using the joint observations of various features seen by the stereo-pair cameras, an optimization of camera parameters (e.g., orientation, distortion, focal length, etc.) and object point coordinates was performed. The two GNSS stations in the FOV as well as solid rock areas in the background were used as ground control points, defining the geodetic datum. A parameter significance analysis was performed to set appropriate camera parameter weights for the respective pseudo-observation equations: high order distortion coefficients (determined in an initial, independent camera calibration) were given strong weights and all but the camera orientation parameters were given weak weights. Points of interest are in distances between 50 and 220 m from the cameras. Combined with a dominant east-facing direction for the far distant points, the area of high precision was found to be relatively small: within a zone of about 80 m $\times$ 80 m, located at a mean distance of 80 m from the cameras, an average precision ($\sigma$) of 6, 5, and 2 cm was achieved for the East, North, and Height components, respectively.

Daily position solutions of the two GNSS stations in the FOV were integrated into the bundle adjustment by means of accurate ground control points with strong weights. Respective image coordinates of the antenna centers were determined by LSM using six templates for every station and view. The agreement of the estimated image antenna coordinates was found to be 0.3 pixel in average.

The effect of integrating the available GNSS position solutions into the bundle adjustment was studied. Constraining the camera position using the respective GNSS position parameters was found to have the most severe influence on the predicted rock boulder precision, i.e., an improvement by a factor of 6 was achieved. Adding the GNSS station located further downslope (station 2) showed an additional improvement by a factor of 2. Adding the upslope GNSS station, on the other hand, was found to have only a minimal improvement (when the camera's positions are fixed). Thus the net effect of integrating at least three GNSS position solutions (two camera positions and one in the FOV) turned into an improvement of precision by a factor of 12. A similar improvement, although not as large, was found for the external reliability of the estimated coordinates.

To test the absolute accuracy of the reconstruction process, the position of the upper GNSS antenna (at a distance of 128 m) was estimated for various epochs, instead of integrating it as a ground control point. Results show that the reconstruction performed within the given confidence levels (a few centimeters). During the summer months of the year 2013, the measurement sequence of the reference GNSS antenna at station 2 was interrupted and thus no high precision ground control point was available for this period. The resulting position estimates for the upstream GNSS station 1 then showed a systematic offset of about 15 cm for the East and 6 cm for the North components (no effect was observed in the Height component). Due to the projection geometry, an error in the definition of the antenna coordinates in the image templates of about 0.5 pixel in image x-direction is already enough to yield an absolute error of 15 cm in the East component. Two scenarios may be considered to explain why the offsets for the East and North coordinates vanish in case station 2 position solutions were available: a first possibility is that a similar offset error was made when the image coordinates of the GNSS antenna for station 2 were defined in the image templates. Because the influence of station 2 on the definition of the geodetic datum is larger than the influence of the pseudo ground control points, the error can be compensated. This, however, would lead to a systematic offset of all position estimates when station 2 was used as ground control point (not visible in station 1 because the same antenna coordinate offset is acting). The alternative argumentation of the problem is that a systematic error in the estimated image coordinates of the initial calibration with additional ground control points has occurred. In this scenario, the estimated coordinates of pseudo ground control points would have a systematic error such that the observed offset is obtained when station 2 does not tie it to the correct orientation.

The analysis above shows that (1) a continuous support of the GNSS station as ground control point is important as otherwise the relatively weak geometrical setup is affected by systematic errors, and (2) the determination of precise image coordinates of the GNSS antenna centers (as well as the calibration plate coordinates) is of critical importance. If one or the other systematic error is present and periods of solutions with and without GNSS station 2 are computed, position offsets are 'turned on and off', thus the resulting velocity estimates for such transition periods show highly erroneous values, not indicated by the estimated formal uncertainties. As small systematic errors may always be present when human interaction is necessary (i.e., to define antenna coordinates in the image) the importance of continuous GNSS time series becomes justified.

The reconstruction process was successfully applied during the summer months between October 2012 and July 2016. Mean annual displacements for the years 2013, 2014, and 2015 are summarized in Table 7.1.

**Table 7.1:** Annual displacements of the permafrost creep above the Grabengufer rock glacier for the years 2013 to 2015. E = East, N = North, H = Height, 3D = total displacement. Units are in $^{m}/_{year}$.

| | E | N | H | 3D |
|---|---|---|---|---|
| 2013 | -0.320 | 0.284 | -0.238 | 0.495 |
| 2014 | -0.374 | 0.340 | -0.313 | 0.597 |
| 2015 | -0.421 | 0.426 | -0.392 | 0.718 |

A good agreement was obtained when comparing velocity estimates derived from the monoplotting procedure with those estimated in the stereo-view approach (i.e., the absolute difference is in the order of 0.1 $^{mm}/_{day}$). When the direction of permafrost creep is assumed to be constant and the scenery is observed with a wide angle lens, the flow component that is projected into the image varies as a function of image position. For example, for a particle moving at a constant rate from the right to the left image edge, with a 3D trajectory being perpendicular to the optical axis of the camera system, the velocity becomes less and less underestimated before the image center and gets more and more underestimated again afterwards. For the geometrical situation at the Grabengufer field site, the direction of the horizontal flow components in the uppermost part (right part in the image) points to a large degree towards the camera. Due to the increased slope steepness, however, the 3D displacement also has a prominent vertical component. The combination therefrom thus partially compensates the wide-angle projection problem. Because the surface displacement is not perfectly perpendicular to the line-of-sight, absolute velocities in the monoplotting procedure are underestimated in general.

#### Collocation

Collocation is a powerful and delicate technique to predict valuable signal from noisy measurements. Given a correlation between adjacent observations, a covariance function can be used to extract a significant stochastic signal. One task of this thesis was to implement a robust procedure to estimate covariance functions that match the stochastic signal of various observation sequences. It was shown that the choice of the deterministic model plays an important role in case of non-continuous observation trajectories. For the time series of estimated feature positions, correlation lengths in the order of 30-40 days were found for points in distances between 50 and 100 m from the cameras. For features located at larger distances a general increase of the correlation length was observed.

A direct estimation of the velocity components was performed also by the collocation principle, again using the estimated correlation between the observations. The derivation for appropriate error propagation w.r.t. position and velocity was given. As noted above, the choice of the covariance function is critical for non-continuous observation series and has an even larger impact on the respective error estimates. Correlations between observations are exclusively described by the covariance function and the ratio between the stochastic signal and the (ideally) uncorrelated noise controls the respective contributions. An iterative collocation was conducted to sequentially detect outliers, resulting in a robust collocation approach. Predicted position estimates showed formal errors for near field objects (<100 m) in the order of 1 cm for East and North, and below 1 cm for the height components. Estimated errors for the respective velocity components show large variations, i.e., between 0.1 and 10 mm/day for the summer months when measurements were possible. With absolute velocities in the range of 1 mm/day, this result shows that an absolute velocity on the mm-per-day level can only be reliably estimated under very good circumstances. Generally, such a high level of precision could not be achieved for the majority of rock boulders that were being tracked.

The principle of the adaptive least-squares collocation procedure was found to be a powerful tool for the interpolation of inhomogeneous velocity fields. Although not explored in detail, results presented for 2-dimensional flow fields showed a prominent extension pattern in the front region of the observed area. The strongest extension was observed in 2015, being in good agreement with the high precision GNSS time series: the GNSS station in the front is located in the area of highest extension, also showing a clear temporal acceleration during the late summer months of 2015. While the upper station also showed an increased velocity for this period, the degree of acceleration was less pronounced. Using the adaptive collocation principle for an analyzes of the areal displacement field, a zone of significant extension could be localized: the center of this region is located about 20 m East of GNSS station located in the front (station 2).

Summarizing the average annual flow behavior of the observed area, one can conclude that the surface motion has experienced an acceleration of about 0.1 $^{m}/_{year}$ between the years 2013 and 2014, and an acceleration of 0.13 $^{m}/_{year}$ between the years 2014 and 2015. While these rates are based on the average displacements of all observed rock boulders in the permafrost creep area, a more detailed analysis of different regions within the observed FOV showed that the front area has accelerated by more than 0.2 $^{m}/_{year}$ between the years 2014 and 2015. The resulting extension occurs, as shown by the velocity strain rate analysis, also in the front area. Although this area is located at the edge of the observed rock boulders, results indicate a detachment, present in the front area of the measured permafrost creep. The origin of the rock fall, that happened in the year 2010 with a volume of about 4300 $m^3$, was located slightly below this area. Therefore, further observations are necessary to reveal if such an event is about to happen again in future.

Concluding this work, the main goal to retrieve continuous high resolution time series of rock boulder coordinates was achieved. It was shown that the combination of high precision GNSS coordinates with a stereo-pair of optical cameras can lead to high accuracy results also in a difficult geometrical configuration and within an environment, where the cameras and most features seen in the images move. The comparison between displacement estimates of a single camera and the stereo solution showed good agreement, whereas the critical aspects of projection and error estimation in case of monoplotting was addressed. Although 3D coordinates were estimated for selected rock boulders, the principles of collocation and adaptive collocation allowed to obtain accurate and continuous velocity fields throughout the measurement period and within the area of interest. The combination of state-off-the-art low-cost equipment and processing algorithms have great potential for the unattended and continuous monitoring of permafrost creep over extended areas and long time scales. The principles and techniques presented in this thesis can easily be adjusted and applied e.g., to monitor glacial flow, for landslide surveying applications, and similar scenarios. The results help to better understand the interaction between these landforms and the (changing) environment.

# 8 Outlook

A short outlook for improvements in this research field is given here. The ranking is chosen according to the priority:

- GNSS position solutions used as accurate ground control points have to be accurately identified in the images. For example the center of the solar panel would be a well defined target for automatic template matching. This, however, is critical, as the relative position between an absolute offset of the antenna center and the center of the solar panel (in this example) changes in case the rock boulder, where the GNSS station is mounted on, rotates. Thus a reliable method, hardware and/or software, is needed.
- The most limiting factor for the permanent monitoring of the permafrost creep in terms of the usable measurement period was the short time window in summer when there was no snow coverage. A possibility to increase the measurement periods is to place artificial markers (e.g., metal sticks with targets being slightly above the rock surface) on various boulders in the area of interest. This would help (1) for an easy and more accurate tracking, (2) for automatically matching the positions between the views even in difficult geometrical conditions, and (3), to allow the estimation of the respective positions also in case of slight snow coverage.
- The adaptive collocation technique should be explored in more detail. This powerful method is very sensitive to the controlling parameters but results may show interesting deformations of the changing flow field. A good network and well distributed measurement locations also allow to explore 2.5-dimensional variations or the strain-stress condition in various regions. Additionally, the method can be extended to include time as another dimension.
- In order to increase precision and field of view coverage, more than two cameras at suitable locations should be used. Also helping the stability for the reconstruction processes, an increased redundancy and a more evenly distributed precision for the various regions on the area of interest can be obtained.
- A couple of improvements in terms of feature matching procedures can be explored. I.e., a multi-photo constraint matching approach, where features are being matched under the collinearity constraint using one or more additional cameras.

# Bibliography

Ackermann, F. (1984), Digital image correlation: Performance and potential applications in photogrammetry, *The Photogrammetric Record*, *11*(64), 429–439, doi:10.1111/j.1477-9730.1984.tb00505.x.

Akca, D. (2013), Photogrammetric monitoring of an artificially generated shallow landslide, *The Photogrammetric Record*, *28*(142), 178–195, doi:10.1111/phor.12016.

Alcamo, J., J. Moreno, B. Nováky, M. Bindi, R. Corobov, R. Devoy, C. Giannakopoulos, E. Martin, J. Olesen, and A. Shvidenko (2007), *Climate Change 2007: Impacts, Adaptation and Vulneraility. Contribution of Working Group II to the Fourth Assessment Report of the Intergovernmental panel on Climate Change*, UK, 541–580, Cambridge University Press.

Arthur, D., and S. Vassilvitskii (2007), K-means++: The advantages of careful seeding, in *Proceedings of the Eighteenth Annual ACM-SIAM Symposium on Discrete Algorithms*, SODA '07, pp. 1027–1035, Society for Industrial and Applied Mathematics, Philadelphia, PA, USA.

Baarda, W. (1967), *Statistical Concepts in Geodesy*, Bd. 2,Nr. 4, Netherlands Geodetic Commions, Publications on Geodesy.

Baltsavias, E. P. (1991), Multiphoto geometrically constrained matching, Ph.D. thesis, ETH Zürich.

Barron, J. L., D. J. Fleet, and S. S. Beauchemin (1994), Performance of optical flow techniques, *International Journal of Computer Vision*, *12*, 43–77, 10.1007/BF01420984.

Barsch, D. (1996), *Rockglaciers: Indicators for the Present and Former Geoecology in High Mountain Environments*, Springer Series in Physical Environment.

Baumberg, A. (2000), Reliable feature matching across widely separated views, in *Computer Vision and Pattern Recognition, 2000. Proceedings. IEEE Conference on*, vol. 1, pp. 774–781 vol.1, doi:10.1109/CVPR.2000.855899.

Bay, H., A. Ess, T. Tuytelaars, and L. Van Gool (2008), Speeded-up robust features (SURF), *Comput. Vis. Image Underst.*, *110*(3), 346–359, doi:10.1016/j.cviu.2007.09.014.

Bell, R., H. Petschko, M. Röhrs, and A. Dix (2012), Assessment of landslide age, landslide persistence and human impact using airborne laser scanning digital terrain models, *Geografiska Annaler: Series A, Physical Geography*, *94*(1), 135–156, doi:10.1111/j.1468-0459.2012.00454.x.

Berger, M. (1999), Deformable area-based template matching with application to low contrast imagery, Ph.D. thesis, ETH Zürich.

Bernard, É., J. Friedt, F. Tolle, M. Griselin, G. Martin, D. Laffly, and C. Marlin (2013), Monitoring seasonal snow dynamics using ground based high resolution photography (Austre Lovénbreen, Svalbard, 79°N), *ISPRS Journal of Photogrammetry and Remote Sensing*, *75*(0), 92–100, doi:http://dx.doi.org/10.1016/j.isprsjprs.2012.11.001.

Berthier, E., H. Vadon, D. Baratoux, Y. Arnaud, C. Vincent, K. Feigl, F. Rémy, and B. Legrésy (2005), Surface motion of mountain glaciers derived from satellite optical imagery, *Remote Sensing of Environment*, *95*(1), 14 – 28, doi: http://dx.doi.org/10.1016/j.rse.2004.11.005.

Beutel, J., S. Gruber, A. Hasler, R. Lim, A. Meier, C. Plessl, I. Talzi, L. Thiele, C. Tschudin, M. Woehrle, and M. Yuecel (2009), PermaDAQ: A scientific instrument for precision sensing and data recovery under extreme conditions, in *Proceedings of the 8th International Conference on Information Processing in Sensor Networks (IPSN)*, pp. 265–276, ACM.

Beutel, J., B. Buchli, F. Ferrari, M. Keller, L. Thiele, and M. Zimmerling (2011), X-SENSE: sensing in extreme environments., in *Proceedings of Design, Automation and Test in Europe*, pp. 1460–1465, IEEE.

Beutler, G., H. Bock, R. Dach, P. Fridez, A. Gäde, U. Hugentobler, A. Jäggi, M. Meindl, L. Mervart, L. Prange, S. Schaer, T. Springer, C. Urchl, and P. Walser (2007), *Bernese GPS Software Version 5.0*, Astronomical Institute Bern.

Beyer, H. A. (1992), Accurate calibration of CCD-cameras, in *Computer Vision and Pattern Recognition, 1992. Proceedings CVPR '92., 1992 IEEE Computer Society Conference on*, pp. 96–101, doi:10.1109/CVPR.1992.223221.

Blaikie, P., T. Cannon, I. Davis, and B. Wisner (2004), *At Risk: Natural Hazards, People's Vulnerability, and Disasters*, second edition ed., United Kingdom: Routledge.

Brown, D. C. (1971), Close-range camera calibration, *Photogrammetric Engineering*, *37*(8), 855–866.

Brown, D. C. (1976), The bundle adjustment - progress and prospects, in *Int. Arch. Phot., Komm. III*, vol. XXI Part 3, ISP Kongress, Helsinki, invited Paper.

Brown, M. Z., D. Burschka, and G. D. Hager (2003), Advances in computational stereo, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *25*(8), 993–1008, doi:10.1109/TPAMI.2003.1217603.

Buchli, B., F. Sutton, and J. Beutel (2012), GPS-Equipped wireless sensor network node for high-accuracy positioning applications, in *Wireless Sensor Networks: 9th European Conference, EWSN 2012, Trento, Italy, February 15–17, 2012. Proceedings*, edited by G. P. Picco and W. Heinzelman, pp. 179–195, Springer Berlin Heidelberg.

Buchli, B., F. Sutton, J. Beutel, and L. Thiele (2014), Towards enabling uninterrupted long–term operation of solar energy harvesting embedded systems, in *Wireless Sensor Networks: 11th European Conference, EWSN 2014, Oxford, UK, February 17-19, 2014, Proceedings*, edited by B. Krishnamachari, A. L. Murphy, and N. Trigoni, pp. 66–83, Springer International Publishing, Cham.

Bühler, Y., and C. Graf (2013), Sediment transfer mapping in a high-alpine catchment using airborne LiDAR, in *Graf, C. (Red.) Mattertal- ein Tal in Bewegung*, pp. 113–124, Cambridge University Press.

Cardozo, N., and R. Allmendinger (2009), SSPX: A program to compute strain from displacement/velocity data, *Computers & Geosciences*, *35*(6), 1343 – 1357, doi: http://dx.doi.org/10.1016/j.cageo.2008.05.008.

Carosio, A. (2001), The methods of the robust statistics for applications in geodesy and GIS, *Tech. rep.*, Institute for Geodesy and Photogrammetry, ETH Zürich.

Carosio, A. (2008), Fehlertherorie und Ausgleichsrechnung, *Tech. rep.*, Institute for Geodesy and Photogrammetry, ETH Zürich.

Christensen, J., B. Hewitson, A. Busuioc, A. Chen, X. Gao, I. held, R. Jones, R. Kolli, W.-T. Kwon, V. Magaña Rueda, L. Mearns, C. Menéndez, J. Räisänen, A. Rinke, A. Sarr, and P. Whetton (2007), *Regional Climate Projections. In: Climate Change 2007: The Pysical Science Basis. Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change*, Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA.

Clarke, T. A., and J. G. Fryer (1998), The development of camera calibration methods and models, *The Photogrammetric Record*, *16*(91), 51–66, doi:10.1111/0031-868X.00113.

Danuser, G. (1996), Stereo Light Microscope Calibration for 3D Submicron Vision, in *Proceedings of the 18th ISPRS Congress*, vol. XXXI, pp. 101–108.

Danuser, G., and M. Stricker (1998), Parametric model fitting: from inlier characterization to outlier detection, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *20*(3), 263–280, doi:10.1109/34.667884.

De Matías, J., J. J. De Sanjosé, G. López-Nicolás, C. Sagüés, and J. J. Guerrero (2009), Photogrammetric methodology for the production of geomorphologic maps: Application to the veleta rock glacier (sierra nevada, granada, spain), *Remote Sensing*, *1*(4), 829, doi:10.3390/rs1040829.

Debella-Gilo, M., and A. Kääb (2012), Locally adaptive template sizes for matching repeat images of earth surface mass movements, *ISPRS Journal of Photogrammetry and Remote Sensing*, *69*(0), 10–28, doi:http://dx.doi.org/10.1016/j.isprsjprs.2012.02.002.

Delaloye, R., C. Lambiel, R. Lugon, H. Raetzo, and T. Strozzi (2007), Typical ERS InSAR signature of slope movements in a periglacial mountain environment (Swiss Alps), in *Proceedings 'Envisat Symposium 2007'*, Montreux, Switzerland.

Delaloye, R., E. Perruchoud, M. Avian, V. Kaufmann, X. Bodin, H. Hausmann, A. Ikeda, A. Kääb, A. Kellerer-Pirklbauer, K. Krainer, C. Lambiel, D. Mihajlovic, B. Staub, I. Roer, and E. Thibert (2008a), Recent interannual variations of rock glacier creep in the European Alps, in *9th International Conference on Permafrost*, pp. 343–348.

Delaloye, R., T. Strozzi, C. Lambiel, and H. Perruchoud, E. andRaetzo (2008b), Landslide–like development of rockglaciers detected with ERS–1/2 SAR interferometry, in *Proceedings of the FRINGE 2007 Workshop*, Frascati, Italy.

Delaloye, R., C. Lambiel, and I. Gärtner-Roer (2010), Overview of rock glacier kinematics research in the Swiss Alps, *Geographica Helvetica*, *65*(2), 135–145, doi:10.5194/gh-65-135-2010.

Delaloye, R., S. Morard, C. Barboux, D. Abbet, V. Gruber, M. Riedo, and S. Gachet (2013), Rapidly moving rock glaciers in Mattertal, in *Graf, C. (Red.) Mattertal– ein Tal in Bewegung*, pp. 21–31, Cambridge University Press.

Delaloye, R., A. D., C. Barboux, L. Braillard, M. Kummert, and S. Morad (2014), Blockgletscher und Hangrutschungen in Permafrostgebieten, Projekt ›Mattertal‹(2009–2013), *Tech. rep.*, Departement für Geowissenschaften Universität Freiburg, gemeinde St.-Niklaus und Randa, Abschlussbericht 2013.

Dold, J. (1997), Ein hybrides photogrammetrisches Industriemesssystem höchster Genauigkeit und seine überprüfung, Heft 54, Schriftenreihe Studiengang Vermessungswesen, Universität der Bundeswehr, München.

Dumouchel, W., and F. O'Brien (1991), Computing and graphics in statistics, in *Statistical Computing and Statistical Graphics*, edited by A. Buja and P. A. Tukey, chap. Integrating a Robust Option into a Multiple Regression Computing Environment, pp. 41–48, Springer-Verlag New York, Inc., New York, NY, USA.

Egli, R. (2004), Swiss-4D: Geodynamik in der Schweiz, *Tech. rep.*, Swisstopo.

Egli, R., A. Geiger, A. Wiget, and H.-G. Kahle (2007), A modified least-squares collocation method for the determination of crustal deformation: first results in the Swiss Alps, *Geophysical Journal International*, *168*(1), 1–12, doi:10.1111/j.1365-246X.2006.03138.x.

Elmiger, F. C. A. (2002), Geodätische Messtechnik, *Notizen zur Vorlesung Band 2*, ETH Zürich, mit Ergänzungen von B. Bürki.

Farid, H., and E. Simoncelli (2004), Differentiation of discrete multidimensional signals, *Image Processing, IEEE Transactions on*, *13*(4), 496–508, doi: 10.1109/TIP.2004.823819.

Fellbaum, M. (1994), Robuste Bildorientierung in der Nahbereichsphotogrammetrie, Ph.D. thesis, Technischen Universität Braunschweig.

Fermüller, C., D. Shulman, and Y. Aloimonos (2001), The statistics of optical flow, *Computer Vision and Image Understanding*, *82*(1), 1–32, doi: http://dx.doi.org/10.1006/cviu.2000.0900.

Fischler, M. A., and R. C. Bolles (1981), Random sample consensus: A paradigm for model fitting with applications to image analysis and automated cartography, *Commun. ACM*, *24*(6), 381–395, doi:10.1145/358669.358692.

Fisher, P. F., and N. J. Tate (2006), Causes and consequences of error in digital elevation models, *Progress in Physical Geography*, *30*(4), 467–489, doi: 10.1191/0309133306pp492ra.

Fraser, C. S., and M. Shortis (1992), Variation of distortion within the photographic field, *Photogrammetric Engineering and Remote Sensing*, *58*(6), 851–855.

Geiger, A. (1996), Verzerrungs-Analyse, Interpolation und Approximation, *Tech. rep.*, ETH Z̈[u]rich.

Geosystems, L. (2009), *Leica GeoMoS Monitoringsoftware, Product information*, Heerbrugg.

Gruber, S., and W. Haeberli (2007), Permafrost in steep bedrock slopes and its temperature-related destabilization following climate change, *Journal of Geophysical Research: Earth Surface*, *112*(F2), doi:10.1029/2006JF000547, f02S18.

Gruen, A. (1978), Accuraxy, reliability and statistics in close-range photogrammetry, *Proc. of ISP Commission V Symposium, Stockholm, Sweden*, *22*(5), in IAP.

Gruen, A. (1984), Adaptive least squares correlation - concept and first resuls., *Intermediate Research Project Report to Heleva Associates, Inc., Ohio State University, Columbus*, pp. 1–13.

Gruen, A. (1986), Photogrammetrische Punktbestimmung mit der Bündelmethode, in *Mitteilung Nr. 40*, ETH Zürich.

Gruen, A., and D. Akca (2005), Least squares 3D surface and curve matching, *{ISPRS} Journal of Photogrammetry and Remote Sensing*, *59*(3), 151 – 174, doi: http://dx.doi.org/10.1016/j.isprsjprs.2005.02.006.

Gruen, A., and H. Sauermann (1977), Photogrammetric determination of time-dependent variations of details of a glacier surface using a non-metric camera, in *Symposium on dynamics of templerate glaciers and related problems*, pp. 1 – 8.

Gruen, A. W. (1985), Adaptive least squares correlation: A powerful image matching technique, *South African Journal of Photogrammetry, Remote Sensing and Cartography*, *14*, 175–187.

Gubler, S., J. Fiddes, M. Keller, and S. Gruber (2011), Scale-dependent measurement and analysis of ground surface temperature variability in alpine terrain, *The Cryosphere*, *5*(2), 431–443.

Guillaume, S. (2014), Geodetic networks and parameter estimation, *Tech. rep.*, ETH Zürich, Institue of Geodesy and Photogrammetry, Geodesy and Geodynamics Laboratory.

Guillaume, S. (2015), Determination of a precise gravity field for the CLIC feasibility studies, Ph.D. thesis, ETH Zürich.

Guillaume, S., and A. Geiger (2007), Real–time small movement detection with a single GPS carrier phase receiver, in *Proc. ENC-GNSS'07*, pp. 506–516, European Navigation Conference, Geneva, Switzerland.

Häberli, W. (1985), Creep of mountain permafrost: Internal structure and flow of alpine rock glaciers, in *Mitteilungen der Versuchsanstalt für Wasserbau, Hydrologie und Glaziologie*, ETH Zürich, Zürich.

Häming, K., and G. Peters (2010), The structure-from-motion reconstruction pipeline - a survey with focus on short image sequences, *Kybernetika*, *46*(5), 926–937.

Harris, C., and M. Stephens (1988), A combined corner and edge detector, in *Proceedings of the 4th Alvey Vision Conference*, pp. 147–151.

Hartley, R., and F. Schaffalitzky (2004), L infin; minimization in geometric reconstruction problems, in *Computer Vision and Pattern Recognition, 2004. CVPR 2004. Proceedings of the 2004 IEEE Computer Society Conference on*, vol. 1, pp. I–504–I–509 Vol.1, doi: 10.1109/CVPR.2004.1315073.

Hartley, R., and A. Zisserman (2003), *Multiple View Geometry in Computer Vision*, 2 ed., Cambridge University Press, New York, NY, USA.

Hartley, R. I. (1994), An algorithm for self calibration from several views, in *Computer Vision and Pattern Recognition, 1994. Proceedings CVPR '94., 1994 IEEE Computer Society Conference on*, pp. 908–912, doi:10.1109/CVPR.1994.323923.

Hartley, R. I., and P. Sturm (1997), Triangulation, *Computer vision and image understanding*, *68*(2), 146–157.

Heid, T., and A. Kääb (2012), Evaluation of existing image matching methods for deriving glacier surface displacements globally from optical satellite imagery, *Remote Sensing of Environment*, *118*, 339 – 355, doi:http://dx.doi.org/10.1016/j.rse.2011.11.024.

Hollenstein, C., A. Geiger, H.-G. Kahle, and G. Veis (2006), CGPS time–series and trajectories of crustal motion along the West Hellenic Arc, *Geophysical Journal International*, *164*(1), 182–191, doi:10.1111/j.1365-246X.2005.02804.x.

Hollenstein, C., M. D. Müller, A. Geiger, and H.-G. Kahle (2008a), GPS-derived coseismic displacements associated with the 2001 Skyros and 2003 Lefkada earthquakes in Greece, *Bulletin of the Seismological Society of America*, *98*(1), 149–161, doi: 10.1785/0120060259.

Hollenstein, C., M. Müller, A. Geiger, and H.-G. Kahle (2008b), Crustal motion and deformation in Greece from a decade of GPS measurements, 1993–2003, *Tectonophysics*, *449*(1–4), 17–40, doi:10.1016/j.tecto.2007.12.006.

Horn, B. K., and B. G. Schunck (1981), Determining optical flow, *Artificial Intelligence*, *17*(1–3), 185–203, doi:http://dx.doi.org/10.1016/0004-3702(81)90024-2.

Huber, P. J. (1964), Robust estimation of a location parameter, *Annals of Mathematical Statistics*, *35*(1), 73–101, doi:10.1214/aoms/1177703732.

Hurter, F. (2014), GNSS meteorology in spatially dense networks, Ph.D. thesis, ETH Zürich.

Jäckli, H. (1957), Gegenwartsgeologie des bündnerischen Rheingebietes, in *Beiträge zur Geologie der Schweiz*, Geotechnische Serie 36.

James H. McClellan, M. A. Y., Ronald W. Schafer (2003), *Signal Processing First*, Band 1, Pearson/Prentice Hall, iSBN 0130909998, 9780130909992.

Journel, A. G. (1989), Fundamentals of geostatistics in five lessons, *Short Course in Geology, Volume 8*.

Kääb, A. (2002), Monitoring high-mountain terrain deformation from repeated air- and spaceborne optical data: examples using digital aerial imagery and ASTER data, *ISPRS Journal of Photogrammetry and Remote Sensing*, *57*(1–2), 39–52, doi:10.1016/S0924-2716(02)00114-4, geomatics in Mountainous Areas – The International Year of the Mountains, 2002.

Kääb, A., and M. Vollmer (2000), Surface geometry, thickness changes and flow fields on creeping mountain permafrost: Automatic extraction by digital image analysis, *Permafrost and Periglacial Processes*, *11*(4), 315–326, doi:10.1002/1099-1530(200012)11:4<315::AID-PPP365>3.0.CO;2-J.

Kääb, A., W. Haeberli, and G. H. Gudmundsson (1997), Analysing the creep of mountain permafrost using high precision aerial photogrammetry: 25 years of monitoring Gruben rock glacier, Swiss Alps, *Permafrost and Periglacial Processes*, *8*(4), 409–426, doi:10.1002/(SICI)1099-1530(199710/12)8:4<409::AID-PPP267>3.0.CO;2-C.

Kahle, H.-G., M. V. Müller, A. Geiger, G. Danuser, S. Mueller, G. Veis, H. Billiris, and D. Paradissis (1995), The strain field in northwestern Greece and the Ionian Islands: results inferred from GPS measurements, *Tectonophysics*, *249*(1), 41 – 52, doi: http://dx.doi.org/10.1016/0040-1951(95)00042-L.

Kaufmann, V. (1998), Deformation analysis of the Doesen rock glacier (Austrian Alps, Europe)., *Proceedings of the 7th International Conference on Permafrost*, pp. 551–556., yellowknife, NWT, Canada.

Kaufmann, V. (2012), The evolution of rock glacier monitoring using terrestrial photogrammetry: The example of äusseres Hochebenkar rock glacier (Austria), *Austrian Journal of Earth Sciences*, *105/2*, 63–77.

Kaufmann, V., and R. Ladstädter (2003), Quantitative analysis of rock glacier creep by means of digital photogrammetry using multi-temporal aerial photographs: two case studies in the Austrian Alps, *Permafrost. Swets & Zeilinger, Lisse, the Netherlands.*

Kaufmann, V., and G. Seier (2016), Long–term monitoring of glacier change at GÖSSNITZKEES (Austria) using terrestrial photogrammetry, *ISPRS – International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, pp. 495–502, doi:10.5194/isprs-archives-XLI-B8-495-2016.

Keigo Hirakawa, T. W. P. (2005), Adaptive homogeneity-directed demosaicing algorithm, *IEEE Trans Image Process*, *14(3)*, 360–369.

Keller, M., M. Yuecel, and J. Beutel (2009), High resolution imaging for environmental monitoring applications, in *International Snow Science Workshop 2009: Programme and Abstracts*, pp. 197–201, Davos, Switzerland.

Kersten, T., and H. Mass (1995), Photogrammetric 3D point determination for dam monitoring, in *Optical 3-D Measurement Techniques III*, edited by Gruen/Kahmen, pp. 161–168, Wichman Verlag.

Kilburn, C. R., and D. N. Petley (2003), Forecasting giant, catastrophic slope collapse: lessons from Vajont, Northern Italy, *Geomorphology*, *54*(1–2), 21–32, doi:10.1016/S0169-555X(03)00052-7, studies on Large Volume Landslides.

Koch, K. (2012), *Parameter Estimation and Hypothesis Testing in Linear Models*, Springer-Verlag New York, Inc.

Koch, K.-R. (1988), *Parameter Estimation and Hypothesis Testing in Linear Models*, Springer-Verlag New York, Inc., NY, USA.

Kraus, K. (1997), *Photogrammetry - Advanced Methods and Applications*, vol. 2, 4th ed., Dümmler/Bonn.

Kraus, K. (2007), *Photogrammetry*, vol. 1, 2nd ed., de Gruyter.

Krige, D. (1951), A statistical approach to some basic mine valuation problems on the Witwatersrand, *Journal of the Southern African Institute of Mining and Metallurgy*, *52*(6), 119–139, doi:http://dx.doi.org/.

Lai, S.-H. (2000), Robust image matching under partial occlusion and spatially varying illumination change, *Computer Vision and Image Understanding*, *78*(1), 84 – 98, doi: http://dx.doi.org/10.1006/cviu.1999.0829.

Lambiel, C., and R. Delaloye (2004), Contribution of real-time kinematic GPS in the study of creeping mountain permafrost: examples from the western Swiss Alps, *Permafrost and Periglacial Processes*, *15*(3), 229–241, doi:10.1002/ppp.496.

Lambiel, C., R. Delaloye, T. Strozzi, R. Lugon, and H. Raetzo (2008), ERS InSAR for detecting the rock glacier activity, in *Kane, D.L. & K.M. Hinkel, Proceedings of the 9th International Conference on Permafrost*, pp. 1019–1024, Fairbanks, Alaska 2.

Lewis, J. (1995), Fast normalized cross-correlation, *Vision interface*, *10*(1), 120–123.

Limpach, P., and D. Grimm (2009), Rock glacier monitoring with low-cost GPS receiver, in *Abstract Volume 7th Swiss Geoscience Meeting*, pp. 247–248, Neuchatel, Switzerland.

Limpach, P., A. Geiger, J. Beutel, B. Buchli, V. Wirz, and S. Gruber (2011), Permanent monitoring of rock glaciers with low-cost GPS, in *Abstract Volume 9th Swiss Geoscience Meeting*, vol. 9, p. 373, Platform Geosciences, Swiss Academy of Science, SCNAT.

Lowe, D. G. (2004), Distinctive image features from scale-invariant keypoints, *Int. J. Comput. Vision*, *60*(2), 91–110, doi:10.1023/B:VISI.0000029664.99615.94.

Lucas, B. D., and T. Kanade (1981), An iterative image registration technique with an application to stereo vision (IJCAI), in *Proceedings of the 7th International Joint Conference on Artificial Intelligence (IJCAI '81)*, pp. 674–679.

Luhmann, T., S. Robson, S. Kyle, and I. Harley (2006), *Close Range Photogrammetry: Principles, Methods and Applications*, Whittles Publishing.

Lundgren, P. (2009), The geodetic infrastructure required to study and monitor volcanoes, *Geophysical Research Abstracts, EGU Genereal Assembly 2009*, *11*, 6294.

Maas, H. (1996), Automatic DEM generation by multi-image feature based matching, in *In International Archives of Photogrammetry and Remote Sensing, volume XXXI, Part B3*, pp. 484–489.

Maas, H. G., A. Gruen, and D. Papantoniou (1993), Particle tracking velocimetry in three-dimensional flows, *Experiments in Fluids*, *15*(2), 133–146, doi:10.1007/BF00190953.

MacQueen, J. (1967), Some methods for classification and analysis of multivariate observations, in *Proceedings of the Fifth Berkeley Symposium on Mathematical Statistics and Probability, Volume 1: Statistics*, pp. 281–297, University of California Press, Berkeley, Calif.

Matheron, G. (1963), *Principles of Geostatistics*, vol. 58 no. 8 1246–1266, Economic Geology.

Mayer, H. (1999), Automatic object extraction from aerial imagery – a survey focusing on buildings, *Computer Vision and Image Understanding*, *74*(2), 138 – 149, doi: http://dx.doi.org/10.1006/cviu.1999.0750.

Mitiche, A., and I. B. Ayed (2011), *Motion Based Image Segmentation*, pp. 139–160, Springer, Berlin, Heidelberg.

Möller, T., and B. Trumbore (1997), Fast, minimum storage ray-triangle intersection, *J. Graph. Tools*, *2*(1), 21–28, doi:10.1080/10867651.1997.10487468.

Moore, E. H. (1920), On the reciprocal of the general algebraic matrix, *Bulletin of the American Mathematical Society*, *26*, 394–395.

Moritz, H. (1970), Least-squares estimation in physical geodesy, *Veröffentlichungen der Deutschen Geodätischen Kommission.*

Moritz, H. (1973), Least-squares collocation, *Deutsche Geodätische Kommission, Reihe A: Theoretische Geodäsie*, *75*, 1–91.

Morrison, F. (1977), Azimuth-dependent statistics for interpolating geodetic data, *Bulletin géodésique*, *51*(2), 105–118, doi:10.1007/BF02522280.

Moustapha, M., B. Sudret, J.-M. Bourinet, and B. Guillaume (2015), Adaptive kriging reliability-based design optimization of an automotve body structure under crashworthiness constraints, in *ICASP12 – 12th International Conference on Applications of Statistics and Probability in Civil Engineering.*

Müller, M. D. M. (2011), Analysis of long-term GPS observations in greece (1993–2009) and geodynamic implications for the Eastern Mediterranean, Ph.D. thesis, ETH Zürich.

Nelson, R. C., and J. Aloimonos (1989), Obstacle avoidance using flow field divergence, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *11*(10), 1102–1106, doi:10.1109/34.42840.

Noetzli, J., C. Huggel, M. Hoelzle, and W. Haeberli (2006), GIS-based modelling of rock/ice avalanches from alpine permafrost areas, *Computational Geosciences*, *10*, 161 – 178.

Pedersini, F., A. Sarti, and S. Tubaro (1998), Tracking camera calibration in multi-camera sequences through automatic feature detection and matching, in *Signal Processing Conference (EUSIPCO 1998), 9th European*, pp. 1–4.

Pertl, A. (1984), Digital image correlation with the analytical plotter planicomp C–100, *International Archives of Photogrammetry and Remote Sensing*, *25*((3B)), 874–882.

Peter, Y. F. (2000), Present-day crustal dynamics in the Adriatic-Aegean plate boundary zone inferred from continuous GPS-measurements, Ph.D. thesis, ETH Zürich.

Photometrix (2013), Australis calibration software, photogrammetry Software, Photometrix, PO Box 3023, Kew, VIC 3101, Australia.

Press, F., and R. Siever (2001), *Understanding Earth*, Macmillan Higher Education.

Ramanath, R., W. E. Snyder, G. L. Bilbro, and W. A. S. Iii (2002), Demosaicking methods for bayer color arrays, *Journal of Electronic Imaging*, *11*, 306–315.

Remondino, F., and S. El-Hakim (2006), Image-based 3D modelling: A review, *The Photogrammetric Record*, *21*(115), 269–291, doi:10.1111/j.1477-9730.2006.00383.x.

Remondino, F., and C. Fraser (2006), Digital camera calibration methods: considerations and comparisons, in *International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences,*, vol. Vol. XXXVI, edited by ISPRS, Dresden, Germany.

Roer, I., W. Haeberli, M. Avian, V. Kaufmann, R. Delaloye, C. Lambiel, and A. Kääb (2008), *Observations and Considerations on Destabilizing Active Rock Glacier in the European Alps*, pp. 1505–1510, .

Rudin, W. (2011), *Fourier Analysis on Groups*, John Wiley & Sons.

Rüfenacht, D., M. Brown, J. Beutel, and S. Süsstrunk (2014), Temporally Consistent Snow Cover Estimation from Noisy, Irregularly Sampled Measurements, in *Proc. 9th International Conference on Computer Vision Theoryand Applications (VISAPP)*, pp. 275–283, Lisbon, Portugal.

Rummel, R., and K. P. Schwarz (1977), On the nonhomogeneity of the global covariance function, *Bulletin géodésique*, *51*(2), 93–103, doi:10.1007/BF02522279.

Scaioni, M., L. Barazzetti, A. Giussani, M. Previtali, F. Roncoroni, and M. I. Alba (2014), Photogrammetric techniques for monitoring tunnel deformation, *Earth Science Informatics*, *7*(2), 83–95, doi:10.1007/s12145-014-0152-8.

Schätti, I., and M. Rub (2008), Automatische überwachung von kritischer Bahninfrastruktur, *Bahnbau*, *5*, 18–23.

Slabaugh, G. G. (1999), Computing euler angles from a rotation matrix, 39–63 pp.

Solexperts (2009), *GeoMonitor*, GeoMonitor, Product information, Mönchaltorf.

Strozzi, T., A. Kääb, and R. Frauenfelder (2004), Detecting and quantifying mountain permafrost creep from in situ inventory, space-borne radar interferometry and airborne digital photogrammetry, *International Journal of Remote Sensing*, *25*(15), 2919–2931, doi:10.1080/0143116042000192330.

Strozzi, T., U. Wegmüller, E. Perruchoud, . Delaloye, A. Kääb, and C. Ambrosi (2008), Evolution of a deep-seated rock mass movement observed with satellit SAR interferometry, in *Proceedings of the FRINGE 2007 Workshope*, Proceedings of the FRINGE 2007 Workshop.

Strozzi, T., R. Delaloye, H. Raetzo, and U. Wegmüller (2009), Radar interferometric observations of destabilized rockglaciers, *Proceeding of the FRINGE 2009 workshop, Frascati, Italy.*

Strozzi, T., H. Raetzo, U. Wegmüller, J. Papke, R. Caduff, C. L. Werner, and A. Wiesmann (2015), *Satellite and Terrestrial Radar Interferometry for the Measurement of Slope Deformation*, chap. Satellite and Terrestrial Radar Interferometry for the Measurement of Slope Deformation, pp. 161–165, Springer International Publishing, Cham.

Sturm, P., and B. Triggs (1996), *A factorization based algorithm for multi-image projective structure and motion*, pp. 709–720, Springer, Berlin, Heidelberg, doi:doi:10.1007/3-540-61123-1_183.

Travelletti, J., T. Oppikofer, C. Delacourt, J.-P. Malet, and M. Jaboyedoff (2008), Monitoring landslide displacements during a controlled rain experiment using a long-range terrestrial laser scanning (TLS), *Remote Sensing and Spatial Information Sciences*, *37(B5)*, 485–490, international Archives of Photogrammetry.

Travelletti, J., C. Delacourt, P. Allemand, J.-P. Malet, J. Schmittbuhl, R. Toussaint, and M. Bastard (2012), Correlation of multi-temporal ground-based optical images for landslide monitoring: Application, potential and limitations, *ISPRS Journal of Photogrammetry and Remote Sensing*, *70*(0), 39–55, doi: http://dx.doi.org/10.1016/j.isprsjprs.2012.03.007.

Triggs, B., P. F. McLauchlan, R. I. Hartley, and A. W. Fitzgibbon (2000), Bundle adjustment - a modern synthesis, in *Proceedings of the International Workshop on Vision Algorithms: Theory and Practice*, ICCV '99, pp. 298–372, Springer-Verlag, London, UK.

Tuytelaars, T., and L. Van Gool (2004), Matching widely separated views based on affine invariant regions, *International Journal of Computer Vision*, *59*(1), 61–85, doi: 10.1023/B:VISI.0000020671.28016.e8.

Villiger, A. (2014), Improvement of the kinematic model of Switzerland (Swiss 4D II), Ph.D. thesis, ETH Zürich.

Vogel, C., A. Bauder, and K. Schindler (2012), Optical flow for glacier motion estimation, in *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, XXIInd ISPRS Congress, 25 August - 01 September 2012, Melbourne, Australia.*

Wang, H., and M. Brady (1995), Real-time corner detection algorithm for motion estimation, *Image and Vision Computing*, *13*(9), 695 – 703, doi: http://dx.doi.org/10.1016/0262-8856(95)98864-P.

Whillans, I. M., and Y.-H. Tseng (1995), Automatic tracking of crevasses on satellite images, *Cold Regions Science and Technology*, *23*(2), 201–214, doi: http://dx.doi.org/10.1016/0165-232X(94)00009-M.

Wild, U., S. Grünig, and R. Hug (2006), Automatisches GPS-Netz Schweiz (AGNES): Stationsnetz und Positionierungsdienste, *Berichte–Serie zum Aufbau der neuen Landesvermessung der Schweiz 'LV95'*, *19*, 97.

Wirth, B. (1990), Höhensysteme, Schwerepotentiale und Niveauflächen: systematische Untersuchungen zur zukünftigen terrestrischen und GPS-gestützten Höhenbestimmung in der Schweiz, Ph.D. thesis, ETH Zürich.

Wirz, V., J. Beutel, B. Buchli, S. Gruber, and P. Limpach (2013), Temporal characteristics of different cryosphere-related slope movements in high mountains, in *Landslide Science and Practice: Volume 4: Global Environmental Change*, edited by C. Margottini, P. Canuti, and K. Sassa, pp. 383–390, Springer Berlin Heidelberg.

Wirz, V., J. Beutel, S. Gruber, S. Gubler, and R. Purves (2014), Estimating velocity from noisy GPS data for investigating the temporal variability of slope movements, *Natural Hazards and Earth System Sciences*, *14*(9), 2503–2520.

Wirz, V., M. Geertsema, S. Gruber, and R. S. Purves (2016), Temporal variability of diverse mountain permafrost slope movements derived from multi-year daily GPS data, Mattertal, Switzerland, *Landslides*, *13*(1), 67–83, doi:10.1007/s10346-014-0544-3.

Zhang, L. (2005), Automatic digital surface model (DSM) generation from linear array images, Ph.D. thesis, ETH Zürich, report No. 88.

Zhang, L., and A. Gruen (2006), Multi-image matching for DSM generation from IKONOS imagery, *ISPRS Journal of Photogrammetry and Remote Sensing*, *60*(3), 195 – 211, doi: http://dx.doi.org/10.1016/j.isprsjprs.2006.01.001.

Zheltov, S. Y., and A. Sibiryakov (1997), Adaptive subpixel cross-correlation in a point correspondence problem, In: Gruen, A., Kuhbler, O. (eds.) Optical 3D Measurement Techniques.

Zhllin, L., Z. Xu, M. Cen, and X. Ding (2001), Robust surface matching for automated detection of local deformations using least-median-of-squares estimator, *Photogrammetric Engineering & Remote Sensing*, *67*(11), 1283–1292.

Zitová, B., and J. Flusser (2003), Image registration methods: a survey, *Image and Vision Computing*, *21*(11), 977 – 1000, doi:http://dx.doi.org/10.1016/S0262-8856(03)00137-9.

# A Snow Cover Estimates

Snow cover was estimated on the basis of grey-level pixel intensities using image segmentation principles and statistical properties thereof. A set of critical parameters were determined empirically using a set of training images. About 95% of the images were correctly segmented into areas with and without snow throughout the year. Although this procedure is not state-off-the-art (see Section 3.3.3), reasonably good results were obtained for the current scenario. For the snow cover variation throughout the measurement period, presented in Fig. 3.6, the area outlined in Fig. A.1 was used for reference.

**Figure A.1:** Outline of the Area used for the relative snow cover measurement shown in Fig. 3.6.

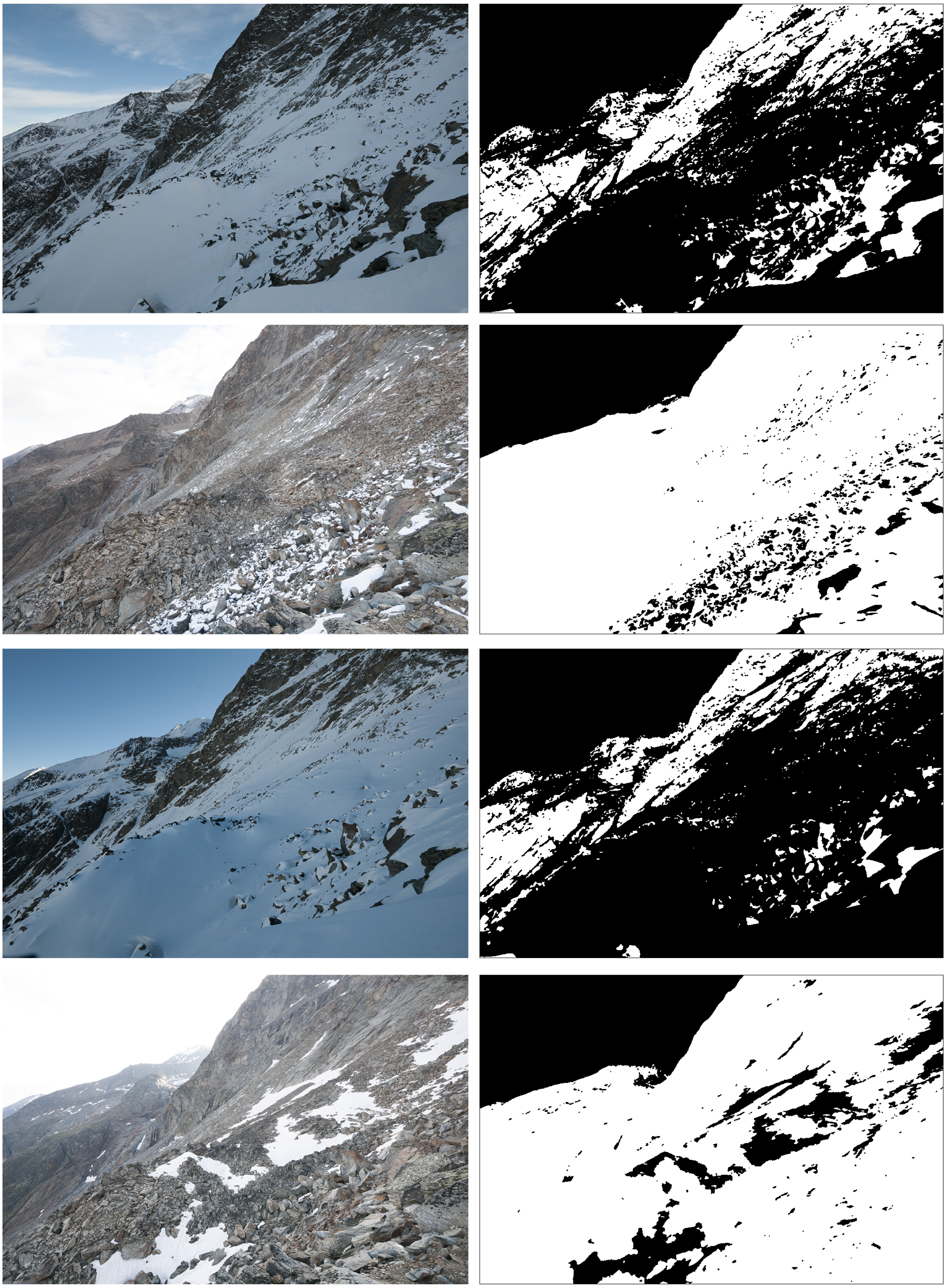

**Figure A.2:** Four examples on detected snow pixels that were mask for the purpose of least-squares feature matching, i.e., Only areas appearing in white (right column of figures) were used.

# B Camera Calibration Results

The following tables summarizes the intrinsic and additional parameters that were determined using the calibration pattern and the Australis software package.

**Table B.1:** Estimated parameters of camera 1 using the Australis camera calibration as self-calibration procedure. A total of 33 and 43 images with 175 points each were used. The $1\sigma$ RMS is 0.32 pixel (1.75 $\mu$).

| Parameter | $\hat{p}$ | $\hat{\sigma}_p$ |
|---|---|---|
| $x_p$ | $0.1583$ | $0.001$ |
| $y_p$ | $-0.0642$ | $0.001$ |
| $c$ | $13.7163$ | $0.001$ |
| $k_1$ | $-1.354e^{-4}$ | $2.133e^{-6}$ |
| $k_2$ | $6.822e^{-7}$ | $3.569e^{-8}$ |
| $k_3$ | $1.026e^{-9}$ | $1.823e^{-10}$ |
| $p_1$ | $-5.362e^{-5}$ | $1.711e^{-6}$ |
| $p_2$ | $-8.024e^{-5}$ | $1.676e^{-6}$ |
| $s_c$ | $0.000$ | $6.604e^{-10}$ |
| $s_h$ | $0.000$ | $6.604e^{-10}$ |

**Table B.2:** Estimated parameters of camera 2 using the Australis camera calibration as self-calibration procedure. A total of 33 and 43 images with 175 points each were used. The $1\sigma$ RMS is 0.4 pixel ($2.21\mu$).

| Parameter | $\hat{p}$ | $\hat{\sigma}_p$ |
|---|---|---|
| $x_p$ | $0.2365$ | $0.002$ |
| $y_p$ | $-0.0620$ | $0.002$ |
| $c$ | $13.8760$ | $0.002$ |
| $k_1$ | $-1.230e^{-4}$ | $2.600e^{-6}$ |
| $k_2$ | $6.608e^{-7}$ | $3.910e^{-8}$ |
| $k_3$ | $0.246e^{-9}$ | $1.788e^{-10}$ |
| $p_1$ | $2.940e^{-5}$ | $2.567e^{-6}$ |
| $p_2$ | $5.104e^{-5}$ | $2.541e^{-6}$ |
| $s_c$ | $0.000$ | $8.457e^{-10}$ |
| $s_h$ | $0.000$ | $8.457e^{-10}$ |

# C Appendix to Object Point Reconstruction

## C.1 Spatial Intersection

For spatial intersection, all camera parameters are treated as observations, $\mathbf{l}$ (each with its corresponding variance component or, if given, with the full variance-covariance matrix). Following the notation and definition given in section 3.2.1, the mathematical model then can be written as:

$$f(\check{\mathbf{l}}, \check{\mathbf{p}}) = f(\mathring{\mathbf{l}} + \delta\hat{\mathbf{l}}, \mathring{\mathbf{p}} + \delta\hat{\mathbf{p}}) + \hat{\mathbf{w}} = \mathbf{0} \tag{C.1}$$

$$\text{with} \quad \mathbf{l} \sim \mathcal{N}(\check{l}; \sigma_0^2 \cdot \mathbf{Q}_{ll}) \tag{C.2}$$

$$\mathbf{p} \sim \mathcal{N}(\check{\mathbf{p}}; \sigma_0^2 \cdot \mathbf{Q}_{pp}) \tag{C.3}$$

More specifically, after linearisation this can be written as:

$$\mathbf{B}^T \cdot \delta\hat{\mathbf{l}} + \mathbf{A} \cdot \delta\hat{\mathbf{p}} + \hat{\mathbf{w}} = 0 \tag{C.4}$$

with $\mathbf{B}^T$ ... coefficient matrix for observations
$\mathbf{A}$ ... coefficient matrix for parameters
$\delta\hat{\mathbf{l}}$ ... increments of observation vector
$\delta\hat{\mathbf{p}}$ ... increments of parameter vector
$\hat{\mathbf{w}}$ ... discrepancy vector

The dimension of $\mathbf{B}^T$ is given by the number of views $v$, i.e., $(2 \cdot v) \times (16 \cdot v + 2 \cdot v) = r \times n$: Each camera view adds two observation equations $2 \cdot v$ and there are 16 camera parameters involved for each view $(16 \cdot v)$ and finally for each view there are two image point measurements $(2 \cdot v)$. The parameter coefficient matrix $\mathbf{A}$ has the dimension $r \times 3$ for every object point:

$$\mathbf{B}^T = \begin{pmatrix} \frac{\partial f_1}{\partial l_1} & \cdots & \frac{\partial f_1}{\partial l_n} \\ \vdots & \ddots & \vdots \\ \frac{\partial f_r}{\partial l_1} & \cdots & \frac{\partial f_r}{\partial l_n} \end{pmatrix} \qquad \mathbf{A} = \begin{pmatrix} \frac{\partial f_1}{\partial p_1} & \frac{\partial f_1}{\partial p_2} & \frac{\partial f_1}{\partial p_3} \\ \vdots & \vdots & \vdots \\ \frac{\partial f_r}{\partial l_1} & \frac{\partial f_r}{\partial p_2} & \frac{\partial f_r}{\partial p_3} \end{pmatrix} \tag{C.5}$$

As indicated in equation C.4, a priori values for all the parameters need to be given in order to start the optimization process. For the

## C.2 Camera Parameter Optimization

To retrieve the camera parameters from a given set of image and object point measurements, the mathematical model is formulated as shown in Section C.1, whereas the observation vector $\mathbf{l}$ combine the image and object points and the unknown parameter vector $\mathbf{p}$ contains camera parameters. To make this parameter estimation procedure more flexible to known parameters or object coordinates (e.g., known camera position or object points being control points), the formulation given in equation C.4 is extended for hard and soft constraints on the parameters and on the observations:

$$\begin{aligned} \mathbf{B}^T \cdot \delta\hat{\mathbf{l}} + \mathbf{A} \cdot \delta\hat{\mathbf{p}} &= -\mathbf{w}_l \\ \mathbf{A}^T \cdot \delta\hat{\mathbf{p}} + \mathbf{I} \cdot \delta\hat{\mathbf{p}} &= -\mathbf{w}_p \end{aligned} \tag{C.6}$$

here $\mathbf{w_l}$ and $\mathbf{w_l}$ are the discrepancy vectors of the true observations ($f(\mathbf{l}, \mathring{\mathbf{p}}) = \mathbf{w}_l$) and the pseudo observations of the parameters ($f(\mathbf{p}, \mathring{\mathbf{p}}) = \mathbf{p} - \mathring{\mathbf{p}} = \mathbf{w}_p$), respectively. In matrix notation, this becomes:

$$\begin{pmatrix} \mathbf{B}^T & \mathbf{A} \\ \mathbf{0} & \mathbf{I} \end{pmatrix} \cdot \begin{pmatrix} \delta\hat{\mathbf{l}} \\ \delta\hat{\mathbf{p}} \end{pmatrix} + \begin{pmatrix} \mathbf{A} \\ \mathbf{I} \end{pmatrix} \cdot \delta\hat{\mathbf{p}} + \begin{pmatrix} \mathbf{w}_l \\ \mathbf{w}_p \end{pmatrix} = \begin{pmatrix} \mathbf{0} \\ \mathbf{0} \end{pmatrix} \tag{C.7}$$

The stochastic model with added pseudo-observations (indicated by $^*$) is given by the variance-covariance matrices $\mathbf{Q}_{ll}$ and $\mathbf{Q}_{pp}$ such that:

$$\mathbf{Q}_{ll}^* = \begin{pmatrix} \mathbf{Q}_{ll} & \mathbf{0} \\ \mathbf{0} & \mathbf{Q}_{pp} \end{pmatrix} \tag{C.8}$$

The lower the variance components in $\mathbf{Q}_{pp}$, the smaller the residual of the pseudo observation $p - \mathring{p} = 0$. Typically only a subset of the parameters will be constraint but the principle remains the same. If a parameter is supposed to be correct (e.g., the camera position is fixed), it is excluded in the estimation process and treated as an observed parameter with its corresponding variance-covariance components.
A similar principle applies to object points observed as three coordinate measurements. In case an object point is supposed to be error free, its corresponding columns in the $\mathbf{B}^T$ matrix are deleted and the columns and rows of $\mathbf{Q_{ll}}$ are removed. Minimizing the cost function now distributes the increments $\delta\mathbf{l}$ onto the remaining measurements.

## C.3 Camera Parameters used for Testing

**Table C.1:** Intrinsic camera parameters used in the synthetic tests, Section 4.4.7. All quantities are with respect to metric image coordinates.

| Parameter | camera 1 | camera 2 | camera 3 | camera 4 |
|---|---|---|---|---|
| $x_p$ | 0.10 | $-0.05$ | 0.20 | 0.00 |
| $y_p$ | $-0.10$ | 0.05 | 0.00 | $-0.20$ |
| $c$ | 15.0 | 15.0 | 15.0 | 15.0 |
| $k_1$ | $1.0e^{-4}$ | 0.0 | $-1.5e^{-4}$ | 0.0 |
| $k_2$ | 0.0 | $1.0e^{-5}$ | $2.5e^{-6}$ | 0.0 |
| $k_3$ | 0.0 | $-2.0e^{-7}$ | 0.0 | 0.0 |
| $p_1$ | $-1.0e^{-5}$ | 0.0 | 0.0 | $2.5e^{-5}$ |
| $p_2$ | $2.0e^{-5}$ | 0.0 | $3.5e^{-5}$ | $-5.0e^{-4}$ |
| $sc$ | 0.0 | 0.0 | 0.0 | 0.0 |
| $sh$ | 0.0 | 0.0 | 0.0 | 0.0 |

**Table C.2:** Extrinsic camera parameters used in the synthetic tests, section 4.4.7.

| | Parameter | camera 1 | camera 2 | camera 3 | camera 4 |
|---|---|---|---|---|---|
| Config. 1 | $\omega$ $[rad]$ | 1.7 | 1.7 | 1.4 | 1.4 |
| | $\phi$ $[rad]$ | 0.0 | 0.0 | 0.0 | 0.0 |
| | $\kappa$ $[rad]$ | 0.0 | 0.0 | 0.0 | 0.0 |
| | $O_X$ $[m]$ | 0.0 | 0.0 | 0.0 | 0.0 |
| | $O_Y$ $[m]$ | 30.0 | 70.0 | 30.0 | 70.0 |
| | $O_Z$ $[m]$ | 30.0 | 30.0 | 70.0 | 70.0 |
| Config. 2 | $\omega$ $[rad]$ | 1.7 | 1.7 | 1.4 | – |
| | $\phi$ $[rad]$ | 0.0 | 0.0 | 0.0 | – |
| | $\kappa$ $[rad]$ | 0.0 | 0.0 | 0.0 | – |
| | $O_X$ $[m]$ | 0.0 | 0.0 | 0.0 | – |
| | $O_Y$ $[m]$ | 25.0 | 75.0 | 50.0 | – |
| | $O_Z$ $[m]$ | 35.6 | 35.6 | 78.9 | – |
| Config. 3 | $\omega$ $[rad]$ | 1.6 | 1.6 | – | – |
| | $\phi$ $[rad]$ | 0.0 | 0.0 | – | – |
| | $\kappa$ $[rad]$ | 0.0 | 0.0 | – | – |
| | $O_X$ $[m]$ | 0.0 | 0.0 | – | – |
| | $O_Y$ $[m]$ | 25.0 | 75.0 | – | – |
| | $O_Z$ $[m]$ | 50.0 | 50.0 | – | – |

# D Complementary Notes on ALSC

## D.1 Variance-Covariance Matrix of a Strain Field

The strain field computed in the ALSC algorithm is primarily used to deform the metric such that the inhomogeneous signal (Eqn. (5.15a)) can be separated from the homogeneous stochastic signal (Eqn. (5.15b)). In many applications (e.g., *Villiger*, 2014; *Cardozo and Allmendinger*, 2009; *Kahle et al.*, 1995), the strain field is of interest and thus also its error component. Because the strain rate is computed directly from the velocity gradient (Eqn. (5.32), Section 5.3.3), the error of the estimated gradient, $\mathbf{Q}_{\nabla\hat{r}\nabla\hat{r}}$, is derived.

To estimate the error of the strain rate tensor ($\mathbf{Q}_{\nabla s_{i,k}\nabla s_{i,k}}$), the second derivative of all components of $\nabla \mathbf{s}_{i,k}$ (compare Eqn. (5.33)) have to be estimated. For a 3-dimensional problem, with $\mathbf{C}_{s_i' s_i'}(x_a, x_b)$ written as $\mathbf{C}_{x_a x_b}$, $\mathbf{C}_{s_i' s_i'}(y_a, y_b)$ written as $\mathbf{C}_{y_a y_b}$, etc., this is:

$$\nabla^2 \mathbf{C} = \begin{pmatrix} \frac{\partial^2 \mathbf{C}_{x_a x_b}}{\partial x_a \partial x_b} & \frac{\partial^2 \mathbf{C}_{x_a x_b}}{\partial y_a \partial y_b} & \frac{\partial^2 \mathbf{C}_{x_a x_b}}{\partial z_a \partial z_b} \\ \frac{\partial^2 \mathbf{C}_{y_a y_b}}{\partial x_a \partial x_b} & \frac{\partial^2 \mathbf{C}_{y_a y_b}}{\partial y_a \partial y_b} & \frac{\partial^2 \mathbf{C}_{y_a y_b}}{\partial z_a \partial z_b} \\ \frac{\partial^2 \mathbf{C}_{z_a z_b}}{\partial x_a \partial x_b} & \frac{\partial^2 \mathbf{C}_{z_a z_b}}{\partial y_a \partial y_b} & \frac{\partial^2 \mathbf{C}_{z_a z_b}}{\partial z_a \partial z_b} \end{pmatrix} \tag{D.1}$$

The partial derivatives with respect to the coordinate systemen axes ($x$, $y$, and $z$) can not be derived directly due to the deformation of the grid (Section 5.3.4). Thus the second derivative of the first element is:

$$\frac{\partial^2 \mathbf{C}_{x_a x_b}}{\partial x_a \partial x_b} = \frac{\partial}{\partial x_a}\Big(\frac{\partial x_b^{tr}}{\partial x_b}\frac{\partial}{\partial x_b^{tr}} + \frac{\partial y_b^{tr}}{\partial x_b}\frac{\partial}{\partial y_b^{tr}} + \frac{\partial z_b^{tr}}{\partial x_b}\frac{\partial}{\partial z_b^{tr}}\Big) \cdot \mathbf{C}_{x_a x_b} \tag{D.2}$$

Applying the product rule and using the abbreviations (also neglecting subscribts '$a$' and '$b$') $\mathbf{X}_x' = \frac{\partial x^{tr}}{\partial x}$, $\mathbf{Y}_x' = \frac{\partial y^{tr}}{\partial x}$, $\mathbf{Z}_x' = \frac{\partial z^{tr}}{\partial x}$, $\mathbf{X}_{xx}' = \frac{\partial^2 x^{tr}}{\partial x^2}$, $\mathbf{Y}_{xx}' = \frac{\partial^2 y^{tr}}{\partial x^2}$, etc. and $\frac{\partial}{\partial x^{tr}} = \partial_{x'}$, $\frac{\partial}{\partial y^{tr}} = \partial_{y'}$, and so on, this leads to

$$\begin{aligned} \frac{\partial^2}{\partial x^2} = \quad & \mathbf{X}_{xx}'\partial_{x'} + \mathbf{X}_x'\frac{\partial}{\partial x}\partial_{x'} \\ & + \mathbf{Y}_{xx}'\partial_{y'} + \mathbf{Y}_x'\frac{\partial}{\partial x}\partial_{y'} \\ & + \mathbf{Z}_{xx}'\partial_{z'} + \mathbf{Z}_x'\frac{\partial}{\partial x}\partial_{z'} \end{aligned} \tag{D.3}$$

Noting that $\frac{\partial}{\partial x} = \mathbf{X}'_x \partial_{x'} + \mathbf{Y}'_x \partial_{y'} + \mathbf{Z}'_x \partial_{z'}$, Eqn. (D.1) can be expressed as:

$$\begin{aligned}\frac{\partial^2}{\partial x^2} &= \mathbf{X}'_{xx}\partial_{x'} + \mathbf{Y}'_{xx}\partial_{y'} + \mathbf{Z}'_{xx}\partial_{z'} \\ &+ \mathbf{X}'^2_x\partial_{x'x'} + \mathbf{Y}'^2_x\partial_{y'y'} + \mathbf{Z}'^2_x\partial_{z'z'} \\ &+ 2\mathbf{X}'_x\mathbf{Y}'_x\partial_{x'y'} + 2\mathbf{X}'_x\mathbf{Z}'_x\partial_{x'z'} + 2\mathbf{Y}'_x\mathbf{Z}'_x\partial_{y'z'}\end{aligned} \tag{D.4}$$

Applying the same mathematical operations to the second element in Eqn. (D.1) leads to:

$$\begin{aligned}\frac{\partial^2}{\partial y^2} &= \mathbf{X}'_{yy}\partial_{x'} + \mathbf{Y}'_{yy}\partial_{y'} + \mathbf{Z}'_{yy}\partial_{z'} \\ &+ \mathbf{X}'^2_y\partial_{x'x'} + \mathbf{Y}'^2_y\partial_{y'y'} + \mathbf{Z}'^2_y\partial_{z'z'} \\ &+ 2\mathbf{X}'_y\mathbf{Y}'_y\partial_{x'y'} + 2\mathbf{X}'_y\mathbf{Z}'_y\partial_{x'z'} + 2\mathbf{Y}'_y\mathbf{Z}'_y\partial_{y'z'}\end{aligned} \tag{D.5}$$

Therefore, Eqn. (D.1) with $< .. >$ as the scalar product can be written as:

$$\nabla^2\mathbf{C} = \begin{pmatrix} <\mathbf{X}^*, \partial^*\mathbf{C}_x> & <\mathbf{Y}^*, \partial^*\mathbf{C}_x> & <\mathbf{Z}^*, \partial^*\mathbf{C}_x> \\ <\mathbf{X}^*, \partial^*\mathbf{C}_y> & <\mathbf{Y}^*, \partial^*\mathbf{C}_y> & <\mathbf{Z}^*, \partial^*\mathbf{C}_y> \\ <\mathbf{X}^*, \partial^*\mathbf{C}_z> & <\mathbf{Y}^*, \partial^*\mathbf{C}_z> & <\mathbf{Z}^*, \partial^*\mathbf{C}_z> \end{pmatrix} \tag{D.6}$$

with

$$\mathbf{X}^* = \begin{pmatrix}\mathbf{X}'_{xx} & \mathbf{Y}'_{xx} & \mathbf{Z}'_{xx} & \mathbf{X}'^2_x & \mathbf{Y}'^2_x & \mathbf{Z}'^2_x & 2\mathbf{X}'_x\mathbf{Y}'_x & 2\mathbf{X}'_x\mathbf{Z}'_x & 2\mathbf{Y}'_x\mathbf{Z}'_x\partial_{y'z'}\end{pmatrix} \tag{D.7}$$

$$\mathbf{Y}^* = \begin{pmatrix}\mathbf{X}'_{yy} & \mathbf{Y}'_{yy} & \mathbf{Z}'_{yy} & \mathbf{X}'^2_y & \mathbf{Y}'^2_y & \mathbf{Z}'^2_y & 2\mathbf{X}'_y\mathbf{Y}'_y & 2\mathbf{X}'_y\mathbf{Z}'_y & 2\mathbf{Y}'_y\mathbf{Z}'_y\partial_{y'z'}\end{pmatrix} \tag{D.8}$$

$$\mathbf{Z}^* = \begin{pmatrix}\mathbf{X}'_{zz} & \mathbf{Y}'_{zz} & \mathbf{Z}'_{zz} & \mathbf{X}'^2_z & \mathbf{Y}'^2_z & \mathbf{Z}'^2_z & 2\mathbf{X}'_z\mathbf{Y}'_z & 2\mathbf{X}'_z\mathbf{Z}'_z & 2\mathbf{Y}'_z\mathbf{Z}'_z\partial_{y'z'}\end{pmatrix} \tag{D.9}$$

$$\partial^* = \begin{pmatrix}\partial_{x'} & \partial_{y'} & \partial_{z'} & \partial_{x'x'} & \partial_{y'y'} & \partial_{z'z'} & \partial_{x'y'} & \partial_{x'z'} & \partial_{y'z'}\end{pmatrix} \tag{D.10}$$

The same principles applies for a dimension other then three.

Having all terms of the gradient variance-covariance matrix defined ($\mathbf{Q}_{\nabla\hat{\tau}'\nabla\hat{\tau}'}$, Eqn (5.83)), the variance-covariance matrix of the strain rate field ($\mathbf{Q}_{\epsilon'\epsilon'}$), can be obtained by propagation. According to Eqn. (5.33), this is:

$$\mathbf{Q}_{\epsilon'\epsilon'} = \frac{1}{2}(\mathbf{Q}_{\nabla\hat{\tau}'\nabla\hat{\tau}'} + \mathbf{Q}^T_{\nabla\hat{\tau}'\nabla\hat{\tau}'}) \tag{D.11}$$

$\mathbf{Q}_{\epsilon'\epsilon'}$ can also be rotated into the principle axis of the strain rate tensor. According to *Peter* (2000), this is obtained by:

$$\mathbf{Q}_{\epsilon'_p\epsilon'_p} = \frac{\partial\mathbf{R}}{\partial\epsilon'} \cdot \mathbf{Q}_{\epsilon'\epsilon'} \cdot \Big(\frac{\partial\mathbf{R}^T}{\partial\epsilon'}\Big)^T \tag{D.12}$$

where the columns of matrix $\mathbf{R}$ contain the eigenvectors of the strain rate tensor $\epsilon'$.

## D.2 Combining Space and Time

Although the fundamental equations of the ALSC principle (Eqns. (5.21), (5.25), (5.25), (5.28), and (5.29)) are given for a general dimension $s$, there are some special concepts worth mentioning when extending the three dimensional (typically cartesian) space with a 'time' component. Regarding the variance-covariance functions (Eqn. (5.8) - (5.10)), the extension with time is accomplished by any combination of the given models, one for the cartesian space $\mathbf{r}$ and one for time $\mathbf{t}$, for example:

$$\begin{aligned} f(\mathbf{r}, \mathbf{t}) &= \sigma_s^2 \cdot f(\mathbf{r}) \cdot f(\mathbf{t}) \\ &= \sigma_s^2 \cdot \frac{1}{1 + u_1 \cdot |\mathbf{r}_i - \mathbf{r}_j|^{n_1}} \cdot e^{-u_2 \cdot |\mathbf{t}_i - \mathbf{t}_j|^{n_2}} \end{aligned} \tag{D.13}$$

This new coordinate space typically has not the same dimensions as the measurement space, for example, there are velocity measurements (3D, signal dimension) at certain positions and time instances (3D + 1D, position dimension). To compute the strain rate tensor (Eqns. (5.33) and (5.34)) a square matrix with partial derivatives is needed. In the example of 3 signal dimensions and 4 position dimensions, this is accomplished by zero padding:

$$\nabla \mathbf{s}_i = \begin{pmatrix} \nabla s_{i_x} \\ \nabla s_{i_y} \\ \nabla s_{i_z} \\ \nabla s_{i_t} \end{pmatrix} = \begin{pmatrix} \frac{\partial s_{i_x}}{\partial x} & \frac{\partial s_{i_x}}{\partial y} & \frac{\partial s_{i_x}}{\partial z} & \frac{\partial s_{i_x}}{\partial t} \\ \frac{\partial s_{i_y}}{\partial x} & \frac{\partial s_{i_y}}{\partial y} & \frac{\partial s_{i_y}}{\partial z} & \frac{\partial s_{i_y}}{\partial t} \\ \frac{\partial s_{i_z}}{\partial x} & \frac{\partial s_{i_z}}{\partial y} & \frac{\partial s_{i_z}}{\partial z} & \frac{\partial s_{i_z}}{\partial t} \\ 0 & 0 & 0 & 0 \end{pmatrix} \tag{D.14}$$

As seen for example in the Eqns. (5.21) and (5.22), the dilation length $\gamma$ is a scalar value used to define the influence range of the dilation for a given dilation center $\mathbf{r}_0$, for all dimensions. Because the influence range for time and space is not uniquely defined (is $\triangle\mathbf{t}$ of one hour being affected equally as for example $\triangle\mathbf{r}$ of three meters?), a dimensional scaling has to be introduced such that by scaling either space $\gamma_{\mathbf{r}}$ or time $\gamma_{\mathbf{t}}$ component, the scalar value $\gamma_{\mathbf{r}} = \gamma_{\mathbf{t}}$ can be used.

# E GNSS Position Time Series

Given below are the position time series estimated for the two cameras (Fig. E.1) and the two stations on the Grabengufer rock slide (Fig. E.2). A local reference station (Base station, see Section 2.2) was used for the computation of the daily static positioning solutions (implemented and operated by Dr. Philippe Limpach, Mathematical and Physical Geodesy, ETH Zürich).

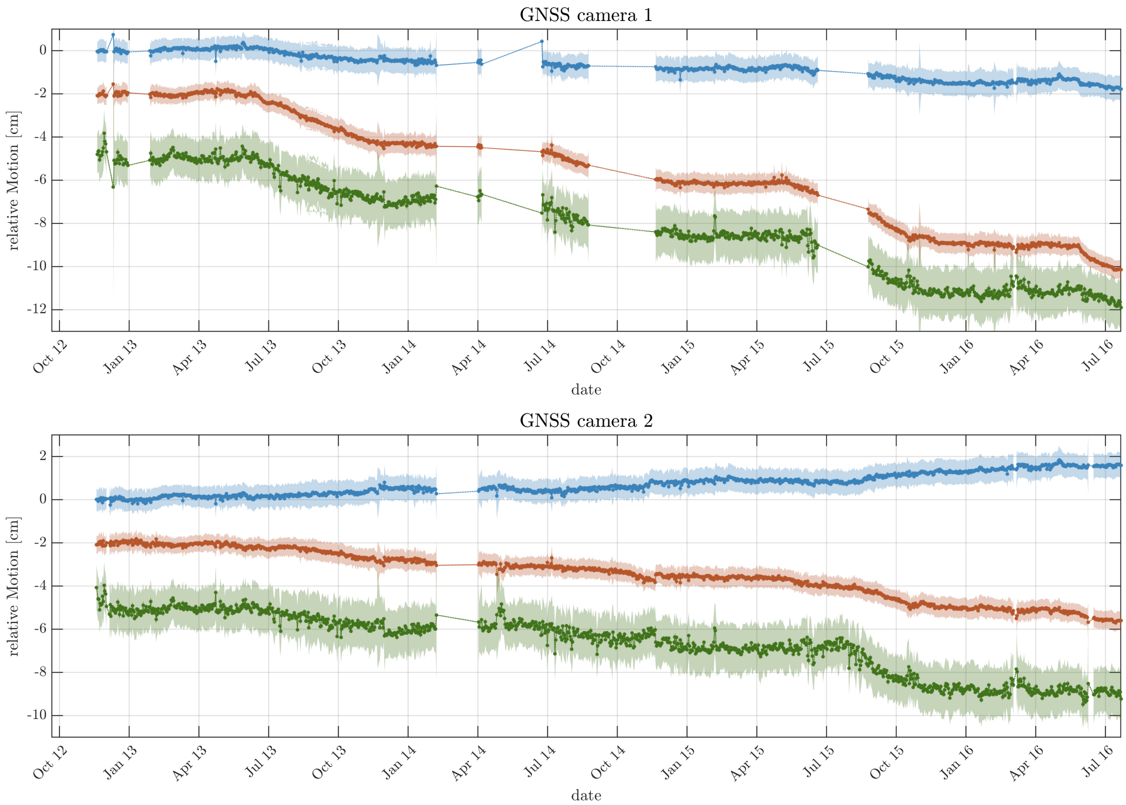


**Figure E.1:** GNSS time series showing the relative camera motion between October 1012 and July 2016. The colored error envelope is the $3\sigma$ level amplified by a factor of 10. Technical issues caused the time series to be not completely continuous. The East, North, and Height components are drawn in red, blue, and green, respectively.

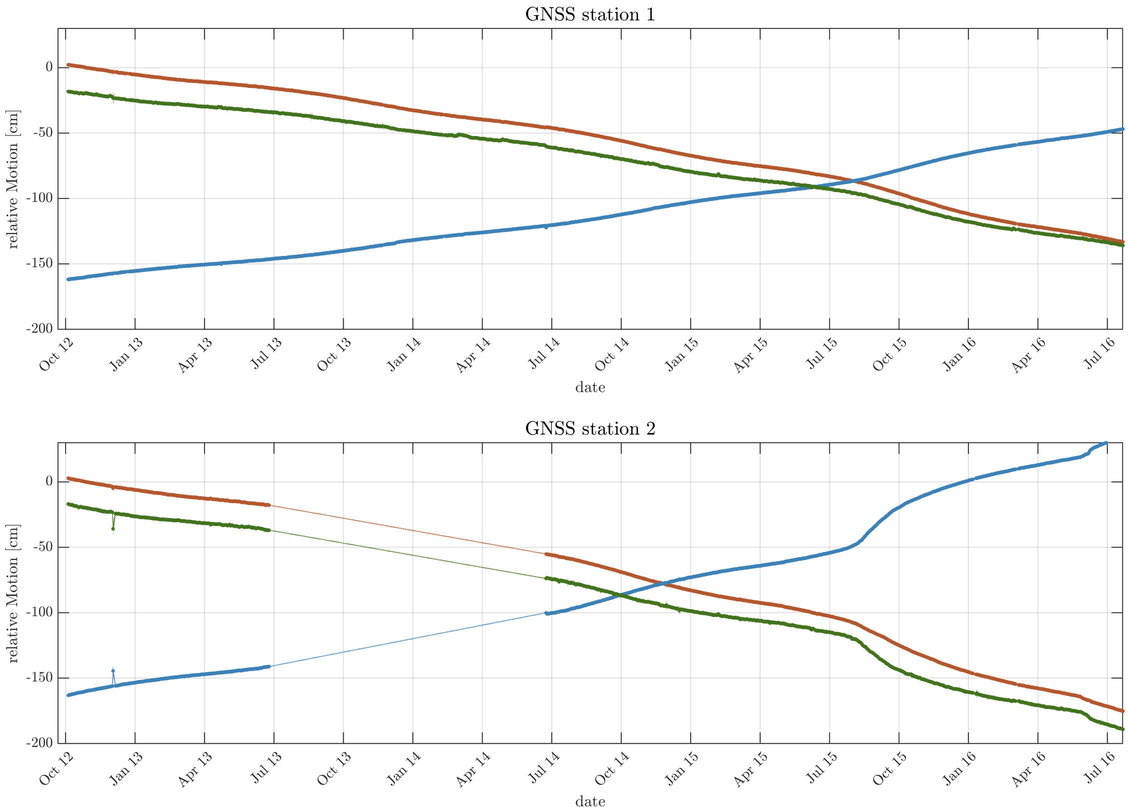


**Figure E.2:** GNSS position time series of stations on the rock glacier. Due to the different scale in the y-axis (compared to Fig. E.1), errors are too small to be shown with the same amplification. GNSS station 2 is located further downslope. The East, North, and Height components are drawn in red, blue, and green, respectively.

For the two stations on the Grabengufer rock slide, i.e., within the FOV of the cameras, the projected motion ($\sim 1 - 2\,{}^{\mathrm{mm}}/_{\mathrm{day}}$ for the total absolute displacements) in the images are in the order of $0.02\,{}^{\mathrm{pixel}}/_{\mathrm{day}}$.