Institut für Geowissenschaften
Mathematisch-Naturwissenschaftliche Fakultät
Universität Potsdam

# Remote Sensing Applications to Earth Surface Processes in the Eastern Central Andes

## DISSERTATION

von

Benjamin J. Purinton

zur Erlangung des akademischen Grades

DOCTOR RERUM NATURALIUM
»Dr. rer. nat.«

in der Wissenschaftsdisziplin Fernerkundung

eingereicht an der
Mathematisch-Naturwissenschaftlichen Fakultät
der Universität Potsdam

Potsdam, October 2019

Published online at the
Institutional Repository of the University of Potsdam:
https://doi.org/10.25932/publishup-44592
https://nbn-resolving.org/urn:nbn:de:kobv:517-opus4-445926

## General Summary

Geomorphology seeks to characterize the forms, rates, and magnitudes of sediment and water transport that sculpt landscapes. This is generally referred to as earth surface processes, which incorporates the influence of biologic (e.g., vegetation), climatic (e.g., rainfall), and tectonic (e.g., mountain uplift) factors in dictating the transport of water and eroded material. In mountains, high relief and steep slopes combine with strong gradients in rainfall and vegetation to create dynamic expressions of earth surface processes. This same rugged topography presents challenges in data collection and process measurement, where traditional techniques involving detailed observations or physical sampling are difficult to apply at the scale of entire catchments. Herein lies the utility of remote sensing. Remote sensing is defined as any measurement that does not disturb the natural environment, typically via acquisition of images in the visible- to radio-wavelength range of the electromagnetic spectrum. Remote sensing is an especially attractive option for measuring earth surface processes, because large areal measurements can be acquired at much lower cost and effort than traditional methods. These measurements cover not only topographic form, but also climatic and environmental metrics, which are all intertwined in the study of earth surface processes. This dissertation uses remote sensing data ranging from handheld camera-based photo surveying to spaceborne satellite observations to measure the expressions, rates, and magnitudes of earth surface processes in high-mountain catchments of the Eastern Central Andes in Northwest Argentina. This work probes the limits and caveats of remote sensing data and techniques applied to geomorphic research questions, and presents important progress at this disciplinary intersection.

## Allgemeine Zusammenfassung

Die Geomorphologie versucht die Art, Geschwindigkeit und Ausmaße des Sediment- und Wassertransports zu charakterisieren welche zur Formung der Landschaften beitragen. Diese werden im Allgemeinen als Erdoberflächenprozesse bezeichnet, welche den Einfluss biologischer (z.B. Vegetation), klimatischer (z.B. Niederschlag) und tektonischer (z.B. Gebirgshebung) Faktoren auf den Transport von Wasser und das erodierte Material beschreiben. Im Hochgebirge entsteht eine dynamische Wechselwirkung zwischen hohen Reliefs und steilen Hängen und infolge dessen starke Regen- und Vegetationsgradienten. Die gleiche raue Topographie stellt wiederum eine Herausforderung bei der Datenerfassung und Prozessmessung dar, da hier herkömmliche Techniken zur detaillierten Beobachtung oder physikalischen Probenahmen im Maßstab ganzer Einzugsgebiete an ihre Grenzen stoßen. Hier zeigt sich der Nutzen der Fernerkundung. Fernerkundung ist definiert als Messung, welche die natürliche Umgebung nicht stört, typischerweise durch Aufnahme von Bildern im sichtbaren bis Radio-Wellenlängenbereich des elektromagnetischen Spektrums. Fernerkundung ist eine besonders vorteilhafte Option für die Messung von Erdoberflächenprozessen, da großflächige Messungen mit wesentlich geringerem Aufwand als bei herkömmlichen Methoden durchgeführt werden können. Diese Messungen ermöglichen nicht nur das Erfassen der topografischen Form, sondern auch das der Klima- und Umwelteinflüsse, die wiederum bei der Untersuchung von Erdoberflächenprozessen miteinander verknüpft sind. In dieser Dissertation werden Fernerkundungsdaten verwendet, die von kamerabasierten Handaufnahmen bis zu weltraumgestützten Satellitenbeobachtungen reichen, um die Auswirkungen, Geschwindigkeiten und das Ausmaß von Erdoberflächenprozessen in hochgebirgigen Einzugsgebieten der östlichen Zentralanden im Nordwesten Argentiniens zu messen. Diese Arbeit untersucht die Möglichkeiten und Grenzen von Fernerkundungsdaten und -techniken, die auf geomorphologische Forschungsfragen angewendet werden und präsentiert wichtige Fortschritte an diesem disziplinären Schnittpunkt.

# Abstract

Remote sensing provides new tools and techniques for earth surface processes research at greater scales and with lower effort than ever before. As remote sensing technology increases and techniques become more refined, geomorphologists are increasingly faced with the potentials, limits, and caveats of new data. Studying earth surface processes at large scales with remote sensing data requires careful consideration of bias and uncertainties in data collection in order to minimize errors in the interpretation of results. In the high mountains, spatial and temporal changes in expressions, magnitudes, and rates of earth surface processes may be more apparent than in low-relief areas, but the uncertainties caused by complex topography also increase.

This dissertation aims to bridge gaps between sparse traditional in situ measurements and observations done at the plot-scale (1–10,000 $m^2$) with dense catchment-scale measurements done through quantitative analysis of remote sensing data in the Eastern Central Andes of Northwest Argentina. The catchments studied in this dissertation straddle the high-elevation, internally drained Altiplano-Puna Plateau, with headwater elevations near 5 km. Elevation decreases rapidly by 4 km over a ~100-km distance from the headwaters in the west to the populated foreland in the east. This high-relief topography creates a steep moisture gradient (< 0.2 to > 1 m/yr precipitation) due to orographic rainfall blocking, which also corresponds to vegetation changes from agricultural fields interspersed with dense subtropical forests in the foreland to succulent-covered, arid hillslopes towards the orogenic plateau. The hillslopes and channel networks in these catchments are prone to high rates and magnitudes of change during extreme rainfall events leading to flooding, channel-bed gravel transport, landslides, and debris flows. These processes are further influenced by anthropogenic mining and construction.

The geomorphic character of this region observed during fieldwork was studied using diverse remote sensing data acquired both on the ground using handheld equipment and from space using satellite sensors. In a first study, this dissertation assesses the ability of global-scale spaceborne elevation dataset differencing for measuring land-level changes in steep topography. The results show that time series of gridded, globally available elevation data are useful for measuring land-level changes, but only when the magnitudes and rates of change exceed expected uncertainties. This constraint limits measurements to lower-slope channel-bed aggradation and incision in regions with high mass-transfer rates influenced by anthropogenic tampering. Steeper, landslide-prone hillslopes are more difficult to measure with relatively low-accuracy globally available spaceborne datasets.

Building on the results from the dynamic aggrading and incising channels in the study area, this dissertation then assesses the limits of camera-based grain-size detection of the channel bedload at the catchment-scale. Previous digital grain-sizing methods focus on lower-relief environments with smaller ranges in bedload grain size and less complex packing arrangements. These techniques provide dissatisfying results in more dynamic, high-mountain rivers. This motivated the development, firstly, of new channel cross-section surveying techniques using the principals of structure-from-motion photogrammetry with a commercial camera attached to a mast. These high-resolution surveys (~1 mm/pixel) were then used to develop new image-segmentation algorithms for grain identification and sizing. Grain segmentation from 3D point clouds was also explored, but this requires much greater effort in photo-survey tactics. The algorithms are published as open-source tools, and the surveying and gravel-sizing techniques have high potential for up-scaling to generate robust grain-size distribution measurements of many field sites in dynamic high-mountain catchments.

The final study in this dissertation also concerns the channel-bedload grain size of high mountain mixed sand- and gravel-bed rivers. Moving from handheld camera-based grain-sizing techniques, this study turns again to spaceborne measurements in the radar wavelength to measure roughness changes of aerially exposed sediment in the channel bed. Although longer wavelength radar-amplitude measurements show greater sensitivity to grain size, this technique is primarily restricted to measuring the transitions between radar-smooth surfaces (sand bars) and radar-rough, gravelly surfaces. These sand-gravel transitions have important implications for studies of downstream fining due to abrasion or selective-transport processes. The use of spaceborne radar data expands the possibility of measuring these transitions to the scale of entire mountain catchments.

Overall, the studies contained in this dissertation probe the limits and caveats of remote sensing data applied to earth surface processes research in the high mountains. The findings show that uncertainties related to sensor biases and complexity of the natural environment must be rigorously assessed prior to interpretation of results. As the spatial resolution, temporal resolution, and accuracy of data increase — and their analysis techniques become more refined — remote sensing will continue to bridge gaps between plot- and catchment-scale measurements in high-mountain environments, towards the goal of synoptic, orogen-scale monitoring of landscape form and rates of change.

## Zusammenfassung

Die Fernerkundung bietet neue Werkzeuge und Techniken zur Analyse von Erdoberflächenprozessen im größeren Umfang und aber geringerem Aufwand als je zuvor. Da die Fernerkundungstechnologie stetig zunimmt und die damit verbundenen Techniken komplexer werden, sind Geomorphologen zunehmend mit den Möglichkeiten, Grenzen und Voraussetzungen neuer Daten konfrontiert. Die Untersuchung von Erdoberflächenprozessen mit Fernerkundungsdaten im großem Maßstab, erfordert Sorgfalt und Einbeziehung von Verzerrungen und Unsicherheiten bei der Datenerfassung. Im Hochgebirge können räumliche und zeitliche Änderungen der Ausprägungen, Größen und Geschwindigkeiten von Erdoberflächenprozessen zwar deutlicher sein als in Gebieten mit niedrigem Relief, dafür werden jedoch die Unsicherheiten durch die komplexe Topographie erhöht.

Diese Dissertation zielt darauf ab, Lücken zwischen spärlichen traditionellen in situ Messungen und Beobachtungen im Parzellenmaßstab (1–10.000 $m^2$), mit Messungen im Einzugsgebietsmaßstab durch quantitative Analyse von Fernerkundungsdaten in den östlichen Zentralanden im Nordwesten Argentiniens zu schließen. Die in dieser Dissertation untersuchten Einzugsgebiete erstrecken sich über das hochgelegene, endorheisch entwässerte Altiplano-Puna Plateau mit Quellgebieten in Höhen von bis zu 5 km. Die Höhe nimmt vom Quellgebiet im Westen bis zum besiedelten Vorland im Osten rasch von 5 km bis zu < 1 km über eine Entfernung von ~100 km ab. Diese hochreliefartige Topographie erzeugt aufgrund der orografischen Niederschlagsbarierre einen steilen Feuchtigkeitsgradienten (< 0,2 bis > 1 m/Jahr Niederschlag). Dies hat auch Vegetationsänderungen von landwirtschaftlichen Feldern zu Folge, die im Vorland von dichten subtropischen Wäldern durchsetzt sind, bis hin zu sukkulenten bedeckten, trockenen Hängen am Fuß des orogenen Plateaus. Die Hänge und Flusssysteme in diesen Einzugsgebieten sind bei extremen Regenfällen hohen Veränderungsraten ausgesetzt, die zu Überschwemmungen, Kiestransporten im Flussbett, Erdrutschen und Trümmerströmen führen. Diese Prozesse werden durch anthropogenen Bergbau und Bau weiter beeinflusst.

Der während der Feldarbeit beobachtete geomorphe Charakter dieser Region wurde mit Hilfe verschiedener Fernerkundungsdaten untersucht, die sowohl am Boden mit Handgeräten als auch aus dem Weltraum mit Satellitensensoren erfasst wurden. In der ersten Studie dieser Dissertation wird die Fähigkeit weltraumgestützter Höhendatensätze zur Messung von Erdoberflächenänderungen innerhalb der steilen Topographie im globalen Maßstab untersucht. Es hat sich gezeigt, dass Zeitreihen von gitterförmigen, global verfügbaren Höhendaten nützlich sind, um Änderungen auf Landebene zu messen, jedoch nur, wenn das Ausmaß und die Raten der Veränderung gößer sind als die zu erwarteten Unsicherheiten. Dies wiederrum beschränkt die Messungen auf Aggradierung und Inzision von Flussbetten mit geringerer Neigung in Regionen mit hohen Massentransportraten, welche durch anthropogene Veränderungen beeinflusst werden. Steilere Hangneigungen, die zu Erdrutschen neigen, sind mit global verfügbaren weltraumgestützten Datensätzen durch deren relativ geringe Genauigkeit schwieriger zu messen.

Aufbauend auf den Ergebnissen der dynamischen Ausweitung und Einschneidung von Flussbetten im Untersuchungsgebiet werden in dieser Dissertation die Grenzen der kamerabasierten Korngrößenerkennung des Geschiebes im Einzugsgebiet untersucht. Bisherige digitale Korngrößenverfahren konzentrieren sich auf Umgebungen mit geringerem Relief, kleineren Korngrößenbereichen und weniger komplexen Packungsanordnungen. In dynamischeren Hochgebirgsflüssen liefern diese Techniken unbefriedigende Ergebnisse . Dies motivierte zum einen die erstmalige Entwicklung neuer Vermessungstechniken für den Flussbettquerschnitt unter Verwendung der Prinzipien der „Structure-from-Motion“ Pho-

togrammetrie unter Verwendung einer kommerziellen Kamera, die an einem Einbeinstativ befestigt wurde. Diese hochauflösenden Datenerhebung (~1 mm/Pixel) wurde dann verwendet, um neue Bildsegmentierungsalgorithmen zur Kornidentifikation und -größenbestimmung zu entwickeln. Die Kornsegmentierung aus 3D-Punktwolken wurde ebenfalls untersucht, was jedoch einen viel größeren Aufwand bei der Fotoerfassungstaktik erfordert. Alle entwickelten Algorithmen werden als Open-Source-Tools veröffentlicht. Die Vermessungs- und Kiesgrößenbestimmungstechniken bieten ein hohes Potenzial für eine Hochskalierung, um robuste Messungen der Korngrößenverteilung vieler Regionen in dynamischen Hochgebirgseinzugsgebieten zu generieren.

Die abschließende Studie befasst sich ebenfalls mit der Korngröße von Flussbetten mit Mischgeschiebe aus Sand und Kies im Hochgebirge. Hier werden im Gegensatz weltraumgestützte Messungen in der Radarwellenlänge benutzt, um Änderungen der Rauheit von an der Oberfläche exponiertem Sediment im Flussbett zu messen. Obwohl länger wellige Radaramplitudenmessungen eine höhere Empfindlichkeit gegenüber der Korngröße zeigen, beschränkt sich diese Technik in erster Linie auf die Messung der Übergänge zwischen radarglatten (Sandbänke) und radarrauen (Kies) Oberflächen. Diese Sand-Kies-Übergänge haben wichtige Implikationen für Untersuchungen der Korngrößenabnahme stromabwärts, welche aufgrund von Abrieb oder selektiven Transportprozessen entsteht. Die Verwendung von weltraumgestützten Radardaten ermöglicht es diese Übergänge auf einem Maßstab ganzer Berggebiete zu übertragen.

Zusammengefasst untersuchen die in dieser Dissertation enthaltenen Studien die Grenzen und Voraussetzungen von Fernerkundungsdaten, die für die Erforschung von Erdoberflächenprozessen im Hochgebirge angewendet werden. Die Ergebnisse zeigen, dass die Unsicherheiten im Zusammenhang mit Sensorungenauigkeiten und die Komplexität der natürlichen Umgebung, vor der Interpretation der Ergebnisse sorgfältig begutachtet werden müssen. Solange die räumliche und zeitliche Auflösung (und Genauigkeit) von Daten zunimmt — und dessen Analysetechniken dadurchkomplexer werden — wird die Fernerkundung weiterhin Lücken zwischen Messungen auf Parzellen- und Einzugsgebietsebene in hochgebirgigen Umgebungen überbrücken müssen, um das Ziel von zusammenfassenden Untersuchung der Landschaftsform und der Änderungsraten im Orogen-Maßstab zu erreichen.

## Acknowledgments

My deep gratitude is extended to my adviser and sole co-author on all three studies in this work, Bodo Bookhagen. Thank you for bringing me along to Berlin five years ago, it altered the course of my life. Santa Barbara and the beach was enticing, but the lakes here are pretty good too. Berlin feels like home, and the life, friends, and partner that I've found here are truly special. I don't know what's ahead, but you, Bodo, have opened me to a world of possibilities, and I thank you.

A big thank you is also extended towards many in the academic realm. Firstly, I thank my co-adviser Manfred Strecker for working tirelessly to create the joint German-Argentine Ph.D. program through which I benefited on many fieldwork trips to the Central Andes over the years. Also, to my mentor Aljoscha Rheinwalt, thank you for providing new perspectives on earth science research and teaching me to always question established methods. Thank you to others in our remote sensing and earth surface processes research group for help throughout, including sitting through my sometimes un-polished presentations and providing helpful suggestions: Stephanie, Simon, Harald, Dominik, Maryam, Mohammad, Ariane, Fabiana, and Bedartha. Thanks also to my field assistants over the years, including Ahmad, Ariane, Pia, and, last but not least, Anna, a true friend in Berlin and a style icon!

A huge thanks go to my close friends, not just in the department, but also in Berlin: Taylor and Magda. (Special thanks to Magda for help with translation of the abstract.) Love you both, thanks for everything over the years! Thanks to my partner, Megan, I love you, too. Your support has been unparalleled, even when you have your own fish to fry. Finally, the greatest love and gratitude go to my parents and sister for all being so supportive of this endeavor over the years and never making me feel guilty about being so far from home, it means a lot to me. Danke schön!

# Contents

**List of Figures** **xiii**

**List of Tables** **xiv**

**1 Dissertation Outline & Author Contributions** **1**
1 Outline . . . 1
2 Publications & Author Contributions . . . 2

**2 Introduction** **3**
1 Contemporary Geomorphology . . . 3
2 Remote Sensing Data For Earth Surface Processes . . . 4
2.1 Near Remote Sensing . . . 4
2.2 Far Remote Sensing . . . 5
3 Study Area . . . 7
4 Motivation and Research Questions . . . 8

**3 Spaceborne DEM Differencing for Land-Level Changes** **11**
1 Introduction . . . 12
2 Spaceborne DEM Errors . . . 13
3 Methods . . . 15
3.1 SRTM-C Correction Steps . . . 16
3.2 Differencing for Change Detection . . . 18
4 Results . . . 19
4.1 Correction Steps . . . 19
4.2 Areas of Change . . . 21
5 Discussion . . . 22
5.1 Necessity of Correction Steps . . . 22
5.2 Potential Change Mapping . . . 27
5.3 Caveats of Data and Method . . . 29
6 Conclusions . . . 30
7 Code and Data Availability . . . 30

**4 Channel-Bed Surveying and Pebble-Counting Algorithms** **33**
1 Introduction . . . 34
1.1 Prior Studies . . . 34
1.2 Motivation . . . 35
2 Algorithm Description . . . 36
2.1 *PebbleCounts*: K-means with Manual Selection (KMS) . . . 37
2.2 *PebbleCountsAuto*: Automatic with Image Filtering (AIF) . . . 39
3 Calibration and Validation I: Controlled Experiment . . . 39
3.1 Experimental Setup . . . 39
3.2 Orthomosaic Generation . . . 39
3.3 Comparison Metrics . . . 40
3.4 Results I: Controlled Experiment . . . 40

4 Calibration and Validation II: Field Surveys . . . 42
4.1 Field Setting . . . 42
4.2 Orthomosaic Generation . . . 42
4.3 Comparison Metrics . . . 44
4.4 Results II: Field Surveys . . . 45
4.5 Caveat of *PebbleCountsAuto* AIF . . . 48
5 Discussion . . . 49
5.1 Performance . . . 49
5.2 Lower Truncation . . . 52
5.3 Image Acquisition . . . 53
5.4 UAV Surveying . . . 53
5.5 Coverage and Processing Limits . . . 54
6 Conclusions . . . 54
7 Code Availability . . . 55

**5 Radar Amplitude Analysis of a Sand- and Gravel-Bed River 57**
1 Introduction . . . 57
2 Study Area and Datasets . . . 59
2.1 Río Toro . . . 59
2.2 SAR Datasets . . . 59
2.3 Use of Radar Over Optical Data . . . 60
3 Methods . . . 61
3.1 SAR and NDVI Data Preprocessing . . . 61
3.2 Surface Roughness Characterization . . . 63
3.3 Backscatter Intensity Analysis . . . 63
4 Results . . . 66
4.1 Endmembers . . . 66
4.2 Field Sites . . . 67
4.3 Downstream Trends . . . 68
4.4 Sand-Gravel Patchiness . . . 69
5 Discussion . . . 70
5.1 Range in Backscatter for Endmembers and Field Sites . . . 70
5.2 Backscatter Trends and Geomorphic Implications . . . 73
5.3 Effect of Polarization . . . 75
5.4 Caveats and Application of the Method . . . 76
6 Conclusions . . . 76

**6 Discussion, Outlook, and Conclusions 79**
1 Research Question 1: Chapter 3 . . . 79
1.1 Future Outlook from Chapter 3 . . . 80
2 Research Question 2: Chapter 4 . . . 81
2.1 Future Outlook from Chapter 4 . . . 81
3 Research Question 3: Chapter 5 . . . 82
3.1 Future Outlook from Chapter 5 . . . 82

4 Concluding Remarks . . . . . 83

**References** **97**

**A Supporting Information for Chapter 3** **98**

**B Supporting Information for Chapter 4** **108**

**C Supporting Information for Chapter 5** **123**

# List of Figures


2.1 Topography, Climate, and Vegetation in the Study Area . . . 9

3.1 Study Area . . . 14
3.2 Co-registration . . . 20
3.3 FFT Destriping . . . 21
3.4 Blocked Shifting . . . 22
3.5 Histograms . . . 23
3.6 QQ-plots . . . 24
3.7 Channel *dh* Profile-view . . . 25
3.8 Channel *dh* Map-view . . . 26
3.9 Landscape *dh* . . . 27

4.1 Plan-view Difference Between KMS and AIF . . . 36
4.2 Distribution Difference Between KMS and AIF . . . 37
4.3 AIF and KMS Flowchart . . . 38
4.4 Lab Results . . . 41
4.5 Decreasing Resolution . . . 42
4.6 Field Sites . . . 43
4.7 Survey Collection . . . 44
4.8 Coarse Binning Results . . . 46
4.9 Measurement Error . . . 47
4.10 Lower Truncation . . . 47
4.11 Truncated Results . . . 49
4.12 Percentile Comparison . . . 50
4.13 High-resolution Results . . . 51
4.14 Percent Match AIF versus KMS . . . 51

5.1 Overview of the Study Area . . . 60
5.2 Characteristics of the Río Toro in Profile . . . 61
5.3 Radar vs. Optical Data . . . 62
5.4 Backscatter from Endmembers . . . 67
5.5 Backscatter from Field Sites . . . 68
5.6 Example Trendline Extraction . . . 70
5.7 Corrected Backscatter Trendlines . . . 71
5.8 Example Power Spectrum Fitting . . . 72
5.9 Power Spectrum Exponents vs. Upstream Distance . . . 73

A1 TanDEM-X Coverage Layer . . . 100
A2 TanDEM-X Height Error Map vs. dGPS . . . 100
A3 TanDEM-X Consistency Map vs. dGPS . . . 101
A4 TanDEM-X Slope Error . . . 102
A5 SRTM-C Slope Error . . . 102
A6 TanDEM-X Height Error Map vs. Slope . . . 103
A7 *dh* vs. Slope and Relief . . . 103

A8 *dh* vs. Elevation . . . . . 104
A9 *dh* vs. Vegetation . . . . . 104
A10 *dh* vs. Curvature . . . . . 105
A11 Aggressive Fourier Filtering . . . . . 106
A12 *dh* vs. Blocked Elevation Adjustment . . . . . 106
A13 *dh* Channel Profile Histogram . . . . . 107

B1 Point-Cloud Density Differences . . . . . 110
B2 Processing Tutorial Part I . . . . . 114
B3 Processing Tutorial Part II . . . . . 115
B4 Percent Matching Grains With Resampling . . . . . 117
B5 Resampling Improvement for S10 . . . . . 117
B6 Resampling Improvement for S34 . . . . . 117
B7 Site-By-Site Comparison of Pebble-Counting Results . . . . . 120
B8 Plan-View of Remaining Issues . . . . . 122

C1 Kernel Density Estimates and Smoothed Trendlines . . . . . 125
C2 TSX/TDX HH Trendline: Single Scene . . . . . 126
C3 S1 VV Trendline: Single Scene . . . . . 127
C4 S1 VH Trendline: Single Scene . . . . . 127
C5 ALOS2 HH Trendline: Single Scene . . . . . 128
C6 ALOS2 HV Trendline: Single Scene . . . . . 128
C7 TSX/TDX HH Trendline: All Scenes . . . . . 129
C8 S1 VV Trendline: All Scenes . . . . . 130
C9 S1 VH Trendline: All Scenes . . . . . 130
C10 ALOS2 HH Trendline: All Scenes . . . . . 131
C11 ALOS2 HV Trendline: All Scenes . . . . . 131
C12 Dry Season Trendlines . . . . . 132
C13 Wet Season Trendlines . . . . . 133
C14 Slope of Power-Law vs. $xmin$ of Fit . . . . . 134

# List of Tables

4.1 *PebbleCounts* vs. Other Algorithms . . . . . 52

B1 Command-Line Variable Flags for *PebbleCounts* . . . . . 112
B2 Command-Line Variable Flags for *PebbleCountsAuto* . . . . . 113

C1 SAR Datasets . . . . . 124

Chapter 1

# Dissertation Outline & Author Contributions

## Outline

This dissertation contains three independent studies on remote sensing applications to earth surface processes in the high-mountain environment of Northwest Argentina. These studies are tied together by an introduction and conclusion, but are otherwise included verbatim as published or submitted in peer-reviewed journals. In Chapter 2, this dissertation provides an introduction to contemporary geomorphology and the application of remote sensing in the field. A description of the study area in the Eastern Central Andes is presented, followed by the motivation for this dissertation and the guiding research questions. The dissertation assesses the research questions in Chapters 3–5, which contain the three articles. Supplementary materials for these chapters are found in the Appendices A–C. In Chapter 6, this dissertation summarizes the progress made in answering the research questions and provides future outlooks and concluding remarks.

# Publications & Author Contributions

The published or submitted journal articles and author contributions are summarized below:

**Chapter 3:** *Purinton, B. and Bookhagen, B.: Measuring decadal vertical land-level changes from SRTM-C (2000) and TanDEM-X (~2015) in the south-central Andes, Earth Surface Dynamics, 6, 971–987, doi: 10.5194/esurf-6-971-2018, 2018.* B. Purinton and B. Bookhagen defined the project to test a variety of spaceborne altimetry measurements for assessing decadal rates and magnitudes of gravel-bed river aggradation and incision, as well as landslide distributions, in the study area. B. Purinton carried out the analysis, developed the Python scripts for SRTM-C correction, produced the figures, and wrote the manuscript. B. Bookhagen provided funding, guidance in data analysis, and manuscript edits.

**Chapter 4:** *Purinton, B. and Bookhagen, B.: Introducing PebbleCounts: a grain-sizing tool for photo surveys of dynamic gravel-bed rivers, Earth Surface Dynamics, 7, 859–877, doi: 10.5194/esurf-7-859-2019, 2019.* B. Purinton and B. Bookhagen defined the project prior to fieldwork in March 2018. B. Purinton devised a camera-on-mast photo-survey setup with input from B. Bookhagen. River-bed cross-sections were surveyed using the setup, and the resulting photo surveys were processed into dense point clouds and seamless orthomosaics by B. Purinton. Grain-size counting algorithms were developed by B. Purinton in Python, with suggestions and guidance from B. Bookhagen. B. Purinton carried out the analysis, developed the Python scripts and manual for *PebbleCounts*, produced the figures, and wrote the manuscript. B. Bookhagen provided funding, guidance in data analysis, and manuscript edits.

**Chapter 5:** *Purinton, B. and Bookhagen, B.: Multiband (X, C, L) radar amplitude analysis for a mixed sand- and gravel-bed river in the Eastern Central Andes, 2019, in review.* B. Purinton and B. Bookhagen defined the project as a novel use of radar backscatter measurements applied to downstream changes in grain size from a large alluvial river with aerially exposed bedload. B. Purinton carried out the analysis, produced the figures, and wrote the manuscript. B. Bookhagen provided funding, guidance in data analysis, and manuscript edits. The study is submitted to the journal *Remote Sensing of Environment* and is in review.

Chapter 2

# Introduction

## Contemporary Geomorphology

Geomorphology is the geoscience subfield concerned with the patterns and forms of landscapes and their evolution through time. In his seminal work on the Western United States, G. K. Gilbert (1877) defined modern geomorphology through detailed descriptions of the competing forces of mountain building, erosion, and transport that generated the predictable shapes of rivers and hillslopes. Over the next 100 years, further contributions were made from natural and lab experiments (e.g., Shields, 1936; Fernandez Luque and Van Beek, 1976), systematic field observations and data collection (e.g., Schumm, 1956; Sundborg, 1956; Hack, 1957), and numerical modeling of landscape evolution (e.g., Ahnert, 1976). These and other studies led to the discovery and formalization of many functional relationships (Dodds and Rothman, 2000) and geomorphic transport laws (Dietrich et al., 2003), through which we now view geomorphology as a predictive science.

Research in geomorphology blossomed at the turn of the 21st century with a number of breakthroughs. These include: innovations in geochemical analysis of cosmogenic radionuclides to infer erosion and soil-production rates (e.g., Bierman and Steig, 1996; Heimsath et al., 1997); theoretical and empirical modeling of river transport and grain-size differentiation processes (e.g., Gomez et al., 2001; Sklar et al., 2006); the development of complex landscape evolution models incorporating the hillslope and fluvial domains at the process level (e.g., Tucker et al., 2001); and the analysis of digital topographic data to determine water flow routing (e.g., Tarboton, 1997), develop hillslope transport laws (e.g., Roering et al., 1999), and constrain the relative influence of tectonic and climatic forcing on shaping rivers and hillslopes (e.g., Burbank et al., 1996; Snyder et al., 2000).

The term *earth surface processes* encapsulates contemporary geomorphic research via the quantitative analysis of multi-disciplinary datasets to understand the mechanisms, rates, and magnitudes of the processes that shape landscapes over sub-annual to million-year time scales. The burgeoning field of remote sensing is responsible for widespread topographic, climatic, and environmental data collection, which has spurred new research and discoveries in the field of earth surface processes at the orogen-scale, using combinations of geochemical dating, topographic analysis, and rainfall and vegetation maps (e.g., Ouimet

et al., 2009; DiBiase et al., 2010; Bookhagen and Strecker, 2012; Scherler et al., 2015; Forte et al., 2016; Olen et al., 2016).

Remote sensing has wide-ranging applications to the study of earth surface processes. This dissertation is concerned with this disciplinary intersection for advancing the use of remote sensing data in geomorphology. In the subsequent sections, types of remote sensing data are elaborated on, an overview of the study area is provided, and the motivation and guiding research questions are presented.

# Remote Sensing Data For Earth Surface Processes

Remote sensing of the earth's surface — or, more broadly, planetary remote sensing — was limited prior to the proliferation of camera-equipped spy airplanes and satellite launches for earth observation in the 1960s and 1970s, driven by Cold War politics. Given this history, remote sensing is often taken to refer to planetary observations from tens to hundreds of kilometers from the surface. But, remote sensing is defined as any measurement that does not disturb the natural environment, typically via active or passive acquisition in the visible- to radio-wavelength range of the electromagnetic spectrum. Remote sensing is an especially attractive option for measuring earth surface processes, as large areal measurements can be acquired at much lower cost and effort than traditional data collection via observation and physical sampling. These measurements cover not only topographic form, but also climatic and environmental metrics, which are intertwined in the study of earth surface processes (e.g., Dietrich and Perron, 2006).

## Near Remote Sensing

The term near remote sensing refers to measurements gathered from the ground (e.g., total station surveying) and aerial systems attached to airplanes, helicopters, and, more recently, unmanned aerial vehicles (UAVs), or drones. A summary of key near remote sensing technology applied to geomorphic research is presented here. This is not a complete review, but rather meant to reflect those technologies with the greatest impact on recent developments in earth surface processes.

### Lidar

Light detection and ranging, or lidar, is an active sensing technology — deployed either on the ground using a Terrestrial Laser Scanner (TLS), or aerially using an Airborne Laser Scanner (ALS) — that records the reflections of a concentrated short-wavelength laser beam to provide point measurements on a target surface. Although this dissertation does not utilize such data, the importance of lidar in earth surface processes research cannot be understated, and is worth briefly mentioning.

Lidar is able to partially penetrate all but the densest vegetation, and can thus provide ground surface models in the form of dense, irregularly spaced point clouds, which are typically gridded to digital terrain models (DTMs) of some desired resolution (Grieve et al., 2016; Smith et al., 2019). Early uses of

lidar-generated high-resolution topography by geomorphologists led to improved transport-law understanding (e.g., Roering et al., 1999, 2007). Since these early days, lidar has become a common tool in earth surface processes research (Roering et al., 2013; Tarolli, 2014; Passalacqua et al., 2015), leading to important developments in river-bed characterization (e.g., Brasington et al., 2012), landscape-scaling relationships (e.g., Clubb et al., 2016; Sangireddy et al., 2017), flow routing (e.g., Rheinwalt et al., 2019), and repeat-pass monitoring for change detection (e.g., Oskin et al., 2012; Mason and Mohrig, 2018).

### Structure-from-Motion and UAVs

The generation of high-resolution topography has become more widespread through Structure-from-Motion with Multiview Stereo (SfM-MVS, or just SfM) photogrammetric developments (Westoby et al., 2012; Fonstad et al., 2013; Smith et al., 2016). SfM relies on the Scale-Invariant Feature Transform (SIFT) algorithm developed by the computer vision community (Snavely et al., 2008) to match distinct features across many images regardless of acquisition geometry and feature distortions. These match points are then used to generate 3D models via the photogrammetric principals of parallax in software packages like *Agisoft Metashape* (Agisoft, 2019). Thus, equipped with a commercial-grade camera and software license, geomorphologists are now able to generate high-resolution topography at significantly lower cost compared to lidar.

The rise of SfM data collection has come in tandem with significant improvements in UAV technology, and the two are increasingly coupled in the geosciences to generate photo surveys and point clouds over large areas (James and Robson, 2014). This union has led to important developments in earth surface processes research including the low-cost monitoring of river-bed aggradation and incision (Cook, 2017), which previously required high-accuracy lidar given the typically small magnitudes of change (e.g., Lane et al., 2003). Additional applications of UAV and SfM technology have been developed in glacial monitoring (e.g., Immerzeel et al., 2014; Rossini et al., 2018) and riparian (e.g., van Iersel et al., 2018) or cryospheric (e.g., Lucieer et al., 2014) vegetation mapping. UAV point clouds and associated orthomosaics have also been used to monitor channel morphology (e.g., Javernick et al., 2014) and bedload characteristics (e.g., Carbonneau et al., 2018), a topic covered extensively in Chapter 4 of this dissertation.

## Far Remote Sensing

The term far remote sensing is reserved here for satellite-based measurements. This has wide applications, from urban sprawl monitoring to natural hazard risk assessment. As such, the focus in this review is again placed on those technologies with the greatest impact on earth surface processes research in recent decades, rather than a full literature review of spaceborne remote sensing.

## Spaceborne DEMs

The United States Geological Survey released the first global (1-km spatial resolution) digital elevation model (DEM) in 1996. Since then, rapid advances in satellite technology have improved the accuracy and increased the resolution of DEMs. The acquisition of the Shuttle Radar Topography Mission (SRTM) 30-m DEM in 2000 (Farr et al., 2007) was the first in a series of near-global $\leq$ 30-m resolution spaceborne DEM missions, continuing up to the current 5-m Advanced Land Observing Satellite (ALOS) World3D (Tadono et al., 2014) and 12-m TanDEM-X DEMs (Rizzoli et al., 2017). The availability of local (e.g., catchment-scale) DEMs has also increased from high-resolution optical satellites (e.g., Pleiades, WorldView), with DEM spatial resolutions under 5 m. Spaceborne DEMs ushered in a new age of quantitative topographic analysis, which has greatly benefited earth surface processes research and solidified the nascent subfield of geomorphometry (Pike et al., 2009). For instance, the SRTM DEM has been widely applied to earth surface processes studies of mountain-belt morphology (e.g., Bookhagen and Strecker, 2012; Olen et al., 2016; Forte et al., 2016).

Of particular interest to earth surface processes research is the measurement of rates and magnitudes of height changes (e.g., river-bed aggradation and incision). However, due to biases and uncertainties in measurements (Purinton and Bookhagen, 2017), the use of spaceborne DEM time series for measuring geodynamic height changes has been largely restricted to the remote cryosphere (e.g., Brun et al., 2017), where the magnitudes of change on snow and ice outweigh the uncertainties in data collection. Exceptions include studies of lava flows (Bagnardi et al., 2016), rock avalanches (Bessette-Kirton et al., 2018), and large coseismic offsets (Barnhart et al., 2019), but the use of spaceborne DEMs to measure land-level changes in ice-free regions remains challenging. This is the subject of Chapter 3 in this dissertation, where the uncertainties and biases inherent in spaceborne DEM generation from various optical or radar sources are also elaborated on.

## Multispectral Sensors

Multispectral satellite sensors from such platforms as Landsat, the Terra/Aqua Moderate Resolution Imaging Spectroradiometer (MODIS), and Sentinel-2 operate in the visible to infrared range, passively measuring the reflectance of the earth's surface and atmosphere. Differences in reflectance at various wavelengths provide global vegetation metrics, such as the Normalized Difference or Enhanced Vegetation Index (NDVI or EVI) (Huete et al., 1994). Additional multispectral sensors operating at longer wavelengths also provide important climatic parameters, such as the measurements of mean annual rainfall or extreme events generated from the Tropical Rainfall Measurement Mission (TRMM) (e.g., Bookhagen and Strecker, 2008; Bookhagen and Burbank, 2010). In conjunction with environmental measurements from sensors like MODIS, these remote data are useful in teasing apart the relative importance of tectonics, climate, and vegetation on landscape form (e.g., Olen et al., 2016; Forte et al., 2016).

### Radar Measurements

A final topic in remote sensing data for earth surface processes is radar-based measurements. Briefly, the use of Interferometric Synthetic Aperture Radar (InSAR) is often used for studies of tectonic deformation (e.g., Bürgmann et al., 2006) but has also been used to measure lava flow subsidence (e.g., Bagnardi et al., 2016). Besides this use of radar phase to measure surface changes, the amplitude of radar backscatter has been applied to measure surface roughness, but this has been restricted to dry desert alluvial fans (Farr and Chadwick, 1996; Hetz et al., 2016) and permafrost landscapes (Regmi et al., 2012). The underexplored use of radar in earth surface processes research is the topic of Chapter 5 in this dissertation.

## Study Area

An overview of the Eastern Central Andes in Northwest Argentina is shown in Figure 2.1. Each chapter contains additional background on this region, but the setting is briefly summarized here. Special attention is given to the strong gradients that create high rates of change, and also to the limited vegetation cover beyond the orographic mountain-front barrier, which together create ideal conditions for remote sensing applied to bare-earth surface processes.

Topographic expression is defined by the development of the Altiplano-Puna Plateau (also referred to as the Central Andean Plateau) beginning in the late-Oligocene and continuing up to the late-Pliocene (Allmendinger et al., 1997; Strecker et al., 2007). The elevation of this orogenic plateau is > 3 km (exceeding 6 km on volcanic edifices) and decreases precipitously towards the eastern foreland (~4 km of relief over a ~100-km distance). Effective moisture blocking at the mountain front created arid conditions in the internally drained plateau interior as early as ~9 Ma (Strecker et al., 2007), with subsequent local generation of internally drained basins at the plateau margins, and flushing of deposited fills during periods of drainage network reconnection and moisture impingement (Sobel et al., 2003; Hilley and Strecker, 2005; Pingel et al., 2013). Modern climate circulation is driven by the South American Summer Monsoon and Low-Level Jet (Vera et al., 2006; Castino et al., 2016b), which delivers moisture from the Amazon basin from north to south along the orographic barrier at the mountain front. Precipitation is further modulated by inter-annual variability in the El Niño Southern Oscillation (Strecker et al., 2007).

The studies contained in this dissertation focus on two steep catchments that straddle the margin of the Altiplano-Puna Plateau, crossing the steep topographic, environmental, and climatic gradients that are characteristic of the region perpendicular to the mountain front. In the Humahuaca and Toro Basins, rainfall increases from < 0.2 m/yr at the arid (sparsely distributed succulents) plateau to > 1 m/yr in the humid (cropland and subtropical forest) foreland. This climate gradient, and the frequent occurrence of extreme rainfall events (Castino et al., 2016b, 2017), leads to high rates of sediment transport in the large

alluvial rivers (Río Toro and Río Grande) that drain the basins, reflected in cosmogenic nuclide erosion rates of up to ~1 mm/yr (Bookhagen and Strecker, 2012).

The rivers each have a significant knickpoint, leading to a break in concavity and oversteepening in narrow gorges downstream of the knickpoint. Both basins contain large Quaternary fill deposits in their upstream reaches, generated during downstream baselevel changes created by local tectonic activity and regional climate variability (Schildgen et al., 2016; Tofelde et al., 2017). The steep topography in the study area is also associated with frequent landslide and debris-flow events (e.g., Savi et al., 2016), causing road damage and even loss of life. The foreland population centers around the cities of Jujuy and Salta are responsible for major anthropogenic modification of the natural river systems, with the presence of numerous gravel-mining operations and flow-retarding weirs. Furthermore, this region is strongly affected by climate change, leading to recent changes in rainfall and flow regimes (Castino et al., 2016a,b, 2017).

# Motivation and Research Questions

As remote sensing technology becomes more widespread and techniques become more refined, geomorphologists are increasingly faced with the potentials, limits, and caveats of new data. Studying earth surface processes at large scales with remote sensing data requires careful consideration of bias and uncertainties in data collection in order to minimize errors in the measurement of processes with small rates and magnitudes. In the high mountains these changes may be more apparent, but the uncertainties caused by high relief are also increased.

This dissertation aims to bridge gaps between sparse traditional in situ measurements and observations done at the plot-scale (1–10,000 $m^2$) with dense catchment-scale measurements done through quantitative analysis of near- and far-remote sensing data in a mountain environment with steep slopes, high relief, and strong gradients in rainfall and vegetation. The dissertation is framed around three research questions:

1. What are the limits of global-scale spaceborne DEM differencing for measuring land-level changes in high-relief terrain?
2. What are the limits of camera-based grain-size detection at the scale of entire catchments in steep and dynamic gravel-bed rivers in the high mountains?
3. What is the potential of spaceborne radar for measuring bedload characteristics at the scale of entire catchments in steep and dynamic gravel-bed rivers in the high mountains?

Chapters 3, 4, and 5 correspond to research questions 1, 2, and 3, respectively.

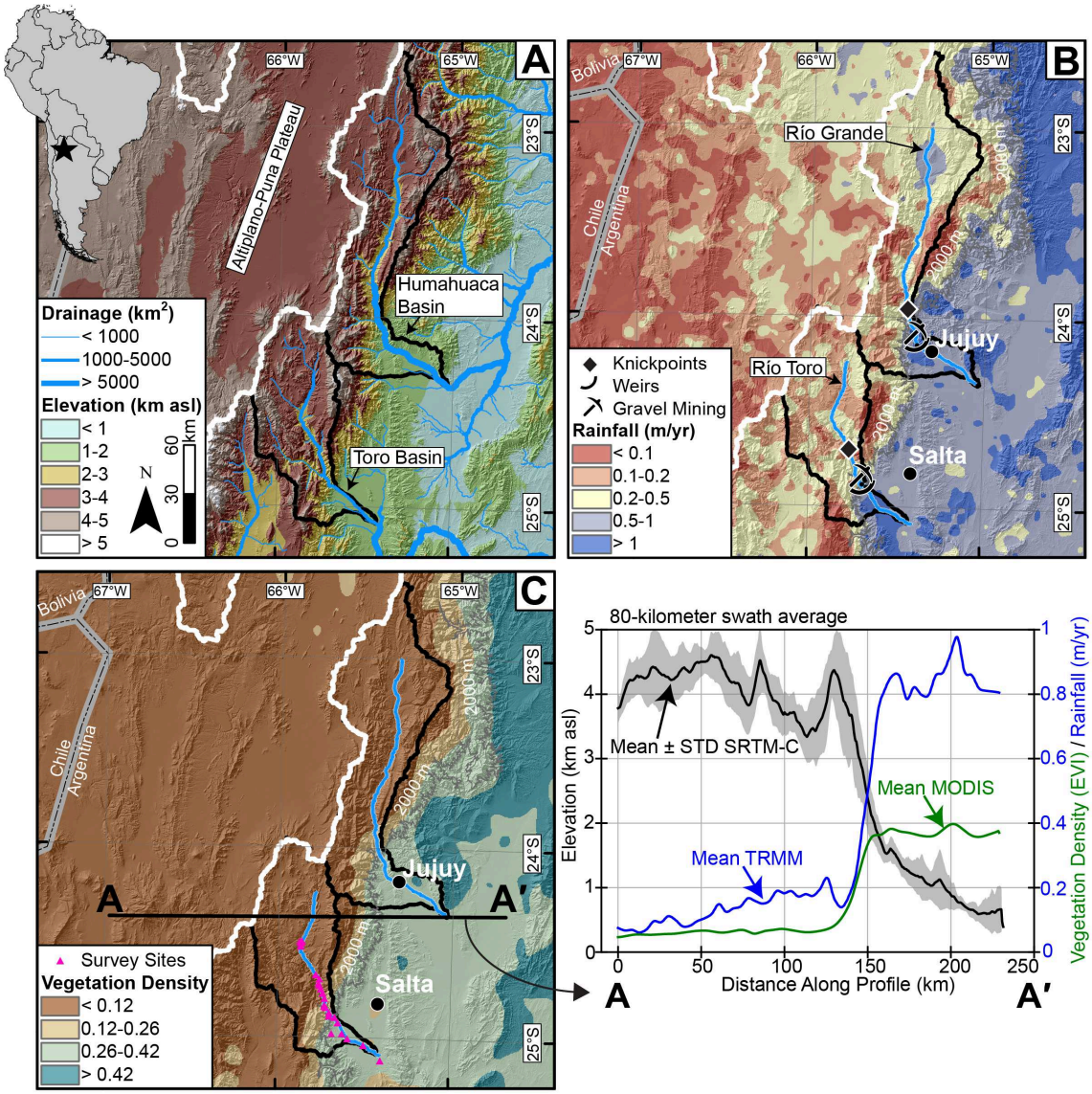


**Figure 2.1** – Overview of study area in NW Argentina with (A) elevation and drainage network (SRTM-C; Jarvis et al. (2008)), (B) rainfall (TRMM2B31 12-year average; Bookhagen and Strecker (2008)), and (C) vegetation (MODIS product 13C1 EVI 14-year average; Huete et al. (1994)), where lower, brown (higher, green) values represent sparse (dense) vegetation. The white watershed boundary delineates the internally drained Altiplano-Puna Plateau. The gray line in (B) and (C) indicates the 2000 m contour line. The strong east-west gradients in all three maps are reflected in the 80-km wide swath profile from A to A′ in (C). Note the anthropogenic tampering of natural gravel-bed channels (Río Grande and Río Toro) with downstream flow diversion (weirs) and gravel-mining activity nearby the populous cities of Salta and Jujuy. The survey sites in (C) refer to the channel-bed photo surveys described in Chapter 4.

Chapter 3

# Measuring decadal vertical land-level changes from SRTM-C (2000) and TanDEM-X (~2015) in the south-central Andes[†]


## Abstract

In the arctic and high mountains it is common to measure vertical changes of ice-sheets and glaciers via digital elevation model (DEM) differencing. This requires the signal of change to outweigh the noise associated with the datasets. Excluding large landslides, on the ice-free earth land-level change is smaller in vertical magnitude and thus requires more accurate DEMs for differencing and identification of change. Previously, this has required meter to sub-meter data at small spatial scales. Following careful corrections, we are able to measure land-level changes in gravel-bed channels and steep hillslopes in the south-central Andes using the SRTM-C (collected in 2000) and the TanDEM-X (collected from 2010–2015) near-global 12–30 m DEMs. Long-standing errors in the SRTM-C are corrected using the TanDEM-X as a control surface and applying cosine-fit co-registration to remove ~1/10 pixel (~3 m) shifts, Fast Fourier Transform and filtering to remove SRTM-C short- and long-wavelength stripes, and blocked shifting to remove remaining complex biases. The datasets are then differenced and outlier pixels are identified as potential signal for the case of gravel-bed channels and hillslopes. We are able to identify signals of incision and aggradation (with magnitudes down to ~3 m in best case) in two > 100 km river reaches, with increased geomorphic activity downstream of knickpoints. Anthropogenic gravel excavation and piling is prominently measured, with magnitudes exceeding $\pm$5 m (up to > 10 m for large piles). These values correspond to conservative average rates of 0.2 to > 0.5 m/yr for vertical changes in gravel-bed rivers. For hillslopes, since we require stricter cutoffs for noise, we are only able to identify one major landslide in the study area with a deposit volume of $16 \pm 0.15 \times 10^6$ m$^3$. Additional signals of change can be garnered from TanDEM-X auxiliary layers, however, these are more difficult to quantify. The methods presented can be extended to any region of the world with SRTM-C and TanDEM-X coverage where vertical land-level changes are of interest, with the caveat that remaining vertical uncertainties in primarily the SRTM-C limit detection in steep and complex topography.


[†] *Purinton, B. and Bookhagen, B.: Measuring decadal vertical land-level changes from SRTM-C (2000) and TanDEM-X (~2015) in the south-central Andes, Earth Surf. Dynam., 6, 971–987, https://doi.org/10.5194/esurf-6-971-2018, 2018.*

# Introduction

Geodynamic and geomorphological processes operating at different time-scales result in vertical change (herein *dh*) on the earth's surface. In the cryosphere, *dh* studies use repeat surveys or digital elevation model (DEM) differencing on annual to sub-annual time-steps (e.g., Berthier et al., 2007; Nuimura et al., 2012; Neelmeijer et al., 2017; Brun et al., 2017). Changes to snow and ice occur most rapidly (aside from landslides), but *dh* measurement outside of the cryosphere also provide aggradation and incision monitoring for rivers (e.g., Lane et al., 2003; Wheaton et al., 2010; Cook, 2017; Mason and Mohrig, 2018), volumes of landslides and extruded lava (e.g., Bagnardi et al., 2016; Bessette-Kirton et al., 2018), and earthquake displacements (Oskin et al., 2012). Large scale monitoring of *dh* on soil, rock, and unconsolidated sediment is an elusive problem requiring signals that outweigh the noise in collection methods and resulting datasets.

Vertical accuracies for modern gridded spaceborne DEMs are on the order of 2–8 m in mountainous regions, though significantly worse on steepening slopes (e.g., Rexer and Hirt, 2014; Purinton and Bookhagen, 2017). Using DEMs from sources like the Advanced Spaceborne Thermal Emission and Reflection Radiometer (ASTER; Tachikawa et al. (2011)) with higher uncertainties is acceptable for monitoring glaciers and ice sheets (e.g., Brun et al., 2017), where *dh* between even sub-annual time-steps can be tens to hundreds of meters over areas of many square kilometers. On the other hand, *dh* of soil, rock, and unconsolidated sediment are often at the centimeter to meter scale and far more localized over up to a few hundred to thousand square meters. Due to these limitations, previous studies relied on intensive mapping from aerial photos (e.g., Hovius et al., 1997), sparse cross-sections with large temporal spans (e.g., Rinaldi and Simon, 1998), or—more recently—meter to sub-meter topographic data from lidar or photogrammetric point clouds (e.g., Lane et al., 2003; Booth et al., 2009; Perroy et al., 2010; Cook, 2017) or select optical satellites with sub-meter resolution like Pleiades and WorldView (e.g., Bagnardi et al., 2016; Bessette-Kirton et al., 2018). Despite recent advances in meter to sub-meter lidar, satellite, and unmanned aerial vehicle data availability (Passalacqua et al., 2015), these remain limited in spatial and temporal coverage, and sometimes prohibitively expensive. Coarser gridded DEMs from radar and optical spaceborne sensors remain the best, and often only, option in large or remote areas.

The publicly available Shuttle Radar Topography Mission (SRTM) DEM is an earth snapshot from its 10 day collection aboard the Endeavour Shuttle in February 2000. The mission produced an Interferometric Synthetic Aperture Radar (InSAR) DEM from C-band (5.6 cm wavelength) radar for 80% of earth's landmasses from typically 2–3 ascending and descending swaths (Farr et al., 2007). The SRTM-C has seen numerous succeeding releases and void filling (e.g., Jarvis et al., 2008). We use the most recent floating point re-processed 1 arcsec (~30 m) NASADEM, taking only the non-void filled original SRTM-C tiles (herein SRTM-C; Crippen et al. (2016); found in the "srtmOnly" directories under: https://e4ftl01.cr.usgs.gov/provisional/MEaSUREs/NASADEM/).

The TanDEM-X 0.4 and 1 arcsec (~12 and ~30 m) DEM released in 2016—here received through scientific DLR proposals, though now available strictly commercially—is the next generation of radar-derived global topography following the SRTM. The TanDEM-X, covering 97% of earth's landmasses, was generated by semi-automated processing and stacking of > 470,000 ascending and descending X-band (3.1 cm wavelength) TerraSAR-X / TanDEM-X satellite bistatic scenes collected from December 2010 to January 2015 (Krieger et al., 2013; Rizzoli et al., 2017).

As elevations are averaged between scenes, we take the date of the TanDEM-X as January 2015, thus providing a 15 year time step of *dh* between SRTM-C and TanDEM-X. Using the latest possible date for TanDEM-X elevations means that rates of change are conservative minimum values.

In this submission we discuss the errors associated with each of these datasets and the corrections applied to mitigate uncertainties in their differencing for *dh* detection outside of the cryosphere. This is therefore a data quality and methods focused study. Geomorphic change detection is applied via correction and differencing of the TanDEM-X and SRTM-C over the south-central Andes in northwestern Argentina (Fig. 3.1) to identify and measure areas of *dh* in gravel-bed channels specifically and then across the landscape. Here, steep gradients in elevation (~1–4 km), rainfall (~0.1–1 m/yr), and vegetation (sub-tropical forests and croplands to arid, succulent-covered slopes) cause high rates of mass transfer (Bookhagen and Strecker, 2012; Savi et al., 2016; Schildgen et al., 2016), further influenced by climate change (Castino et al., 2016a,b, 2017) and anthropogenic modification (gravel mining and weirs). To conclude, we discuss caveats driven by remaining uncertainties prevalent in spaceborne DEMs collected over complex topography.

# Spaceborne DEM Errors

Yamazaki et al. (2017) classify spaceborne DEM errors into speckle noise, stripe noise, absolute bias, and tree height bias. We divide this further for the case of SRTM-C and TanDEM-X (both radar DEMs) into: (i) sensor specific related to radar and spacecraft collection, and (ii) terrain specific related to land-surface cover and topographic complexity. We do not consider DEMs from optical sensors such as ASTER (Tachikawa et al., 2011) and the Advanced Land Observing Satellite (ALOS; Tadono et al. (2014)), which have well documented errors (e.g., Racoviteanu et al., 2007; Nuth and Kääb, 2011; Fisher et al., 2013; Yamazaki et al., 2017) and perform worse than radar, with vertical accuracies > 5 m (1-$\sigma$) and persistent high-frequency artifacts (Purinton and Bookhagen, 2017). Additionally, a dearth of cloud-free, high-quality ASTER imagery covering the study area precludes the automated DEM generation of Girod et al. (2017) and regression techniques of Wang and Kääb (2015). On the other hand, within the study area, the SRTM-C and TanDEM-X both exhibit vertical uncertainties < 3.5 m (Purinton and Bookhagen, 2017) and also have an appropriately long time difference for vertical land-level change detection. Auxiliary rasters including the water indication mask (WAM), height error mask (HEM), consistency mask (COM), and coverage map (COV) delivered with TanDEM-X (Wessel, 2016) allow enhanced understanding of DEM quality (see Appendix A).

Random, or speckle, error caused by instrument thermal noise and localized de-correlation is the primary sensor bias for radar (Rodríguez et al., 2006). These localized, small magnitude errors reduce with increasing looks used in the final mosaic. Speckle presents a greater issue in SRTM-C given the maximum three swaths at lower latitudes (Farr et al., 2007). Such noise is expected to be minimal in the TanDEM-X, with average coverage in our study area of seven ascending and descending scenes, and up to 14 in many steep areas (Fig. A1). Smoothing data prior to and after phase unwrapping (e.g., multi-looking, adaptive filters, or down-sampling) can further reduce speckle. The SRTM-C raw resolution of ~30 m is similar to the final 1 arcsec product, though, due to interferogram smoothing to reduce noise, the estimated true ground resolution of the final product is 45–60 m (Sun et al., 2003; Farr et al.,

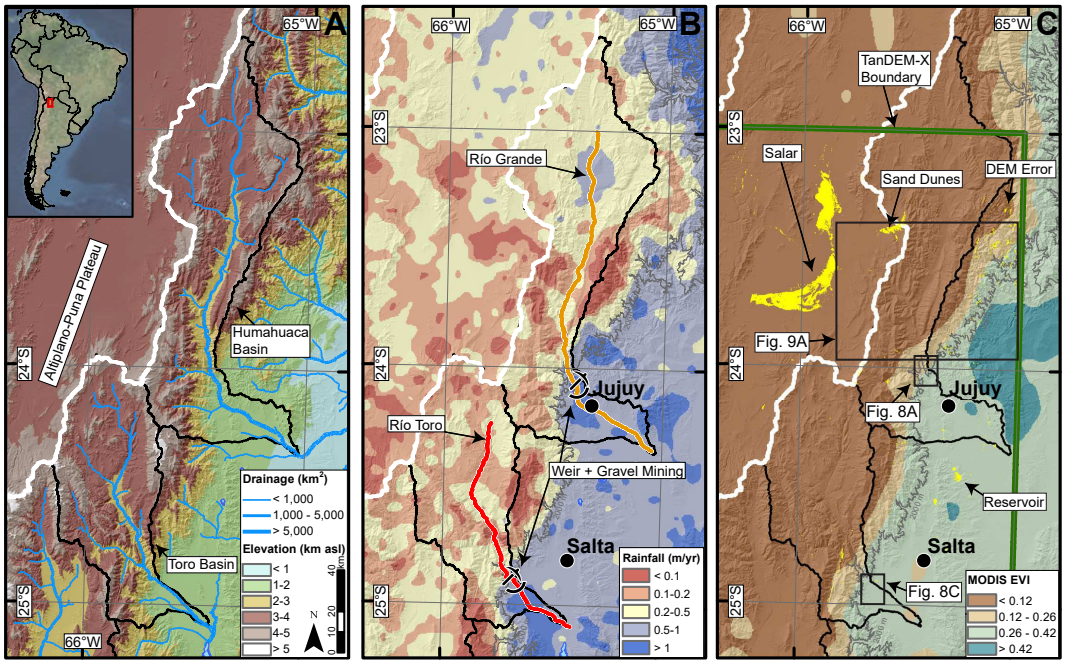


**Figure 3.1** – Overview of study area in NW Argentina with (A) elevation, (B) rainfall (Tropical Rainfall Measurement Mission 12 year average; TRMM2B31; Bookhagen and Strecker (2008)), and (C) vegetation (MODIS product 13C1 Enhanced Vegetation Index 14 year average; MODIS EVI; Huete et al. (1994)), where lower, brown (higher, green) values represent sparse (dense) vegetation. Note strong east-west gradients in all three maps. The white watershed boundary delineates the internally drained Altiplano-Puna Plateau. The gray line in (B) and (C) indicates the 2000 m contour line. The yellow patches in (C) are areas identified in the TanDEM-X water indication mask (WAM) as having low amplitude and/or low coherence. These patches correspond to salt flat (salar) regions on the plateau, water bodies (e.g., reservoirs in the low-elevation areas), steep and vegetated areas (DEM error), and other zones of coherence loss, such as the dunes identified. Inset boxes in (C) indicate locations of *dh* map-view Figures 3.8–3.9, with TanDEM-X tile boundary in green. Note anthropogenic tampering of natural gravel-bed channels (Río Grande and Río Toro) with downstream flow diversion (weirs) and gravel mining activity nearby the populous cities of Salta and Jujuy.

2007; Tachikawa et al., 2011). This may be improved in the newly released data (Crippen et al., 2016), but this remains to be tested. Multi-looking of 4×5 pixels of raw radar returns (resolution ~3.3 m) was used in the case of TanDEM-X to generate a final 0.4 arcsec (~12 m) product, thus significantly smoothing and reducing speckle (Rizzoli et al., 2017).

Besides a small geolocation error expected in both DEMs from instrument uncertainties, the SRTM-C has a number of spacecraft specific biases, manifested in short- and long-wavelength striping (Rodríguez et al., 2006; Yamazaki et al., 2017). The short wavelength (~0.5–1 km, magnitudes typically < 0.5 m) stripes are related to jitter in the antenna mast caused by the periodic firing of shuttle attitude thrusters (Farr et al., 2007). Longer wavelength errors with magnitudes > 1 m are caused by individual swath tilts and form complex undulating patterns over ~100 km distances (Crippen et al., 2016; Yamazaki et al., 2017). TanDEM-X satellite biases can be found in slight tilting of individual TerraSAR-X / TanDEM-X scenes (e.g., Neelmeijer et al., 2017), though these tilts were

removed during stacking in the end product (Rizzoli et al., 2017). The careful monitoring and control maintained over flight geometry, in addition to post-processing to remove tilts using ICESat (Ice, Cloud and land Elevation Satellite; Schutz et al. (2005)), restricts most of the TanDEM-X uncertainty to the second category of terrain specific error (Rizzoli et al., 2017).

Land-surface cover plays a key role in modulating radar returns. TanDEM X-band and SRTM C-band radar have different penetration depths in dense vegetation (Carabajal and Harding, 2006; Hofton et al., 2006; Wessel et al., 2018) and snow and ice (Rignot et al., 2001; Rossi et al., 2016), leading to different height returns. We note this important caveat, but are able to ignore it for our particular study question (land-level change of bare material) and area (only partial vegetation and no permanent snow and ice). Sub-tropical vegetation in our study area does allow some exploration of the effect on *dh*, however, we find no clear relation (see Appendix A). In any case, vegetation differences are expected to be less significant than for optical data, which returns only the canopy heights (e.g., Yamazaki et al., 2017). Both DEMs have major inconsistencies and speckle over water bodies, wet salt flats, and deserts caused by de-correlation, variable reflectance, and/or weak backscatter of the radar signal (Rodríguez et al., 2006; Farr et al., 2007; Wendleder et al., 2013; Rizzoli et al., 2017). For the SRTM-C, these areas are largely voids anyway, and for TanDEM-X the WAM raster provides information on coherence and amplitude for each pixel to identify these untrustworthy measurements (Fig. 3.1C).

Remaining errors in the SRTM-C and TanDEM-X are related to terrain characteristics (see Appendix A). This is the result of topographic complexity below the resolution of the sensor, radar geometry considerations (layover, foreshortening, and shadowing), and interferometric phase unwrapping errors, all most pronounced in steep mountains. Such terrain biases are demonstrated in the SRTM-C with elevation (Berthier et al., 2006; Paul, 2008), slope and aspect (Gorokhovich and Voustianiouk, 2006; Van Niel et al., 2008; Peduzzi et al., 2010; Shortridge and Messina, 2011), and resolution (manifested in curvature) (Gardelle et al., 2012), and in the TanDEM-X with only slope (Purinton and Bookhagen, 2017; Wessel et al., 2018). Terrain slope—also related to relief (Fig. A7)—is the primary cause of error in any DEM, demonstrated in the division of vertical uncertainties for most DEMs into slope bins (e.g., Wessel et al., 2018). Slope dependent errors may be reduced with finer resolution data and increased look angles for mosaicking, as in the case of TanDEM-X, but these uncertainties are expected to remain as the most prevalent cause of error in any spaceborne DEM.

With this framework for understanding the potential error sources in the SRTM-C and TanDEM-X, it is possible to correct one dataset to another in a multi-step processing chain (e.g., Yamazaki et al., 2017) allowing *dh* identification and measurement with greater certainty.

# Methods

Given the excellent agreement with differential GPS globally (Wessel et al., 2018) and in the study area (Purinton and Bookhagen, 2017) along with the minimal errors associated with orbital characteristics, we consider the TanDEM-X DEM as our reference surface in order to correct the more problematic SRTM-C. During correction, we do not apply any speckle reduction (e.g., via an adaptive filter as in Yamazaki et al. (2017)), as we are interested in raw elevation values and not a smoothed DEM. For the SRTM-C we select the non-void filled NASADEM data so as not to include any auxiliary elevation measurements from, for instance, ASTER (Crippen et al., 2016). Importantly,

both DEMs are referenced to the WGS84 ellipsoid vertical datum, whereas previous SRTM-C releases have been referenced to the EGM96 geoid (Farr et al., 2007), thus requiring a geoid-adjustment step introducing additional uncertainties prior to comparison.

For correction and differencing we use the 0.4 arcsec TanDEM-X that we bilinearly resampled to 1 arcsec to match the raw resolution of the SRTM-C. Wessel (2016) note that the delivered TanDEM-X 1 arcsec tiles, which we also have a number of, were generated with average resampling of the 0.4 arcsec tiles by DLR and not by any increase in multi-looks or interferogram smoothing. We tested a number of resampling schemes including average, bilinear, cubic, and cubic spline on the original 0.4 arcsec tiles and found better results (lower vertical uncertainty compared with differential GPS) from the commonly used bilinear resampling, whereas the un-edited 1 arcsec tiles delivered by the DLR—generated by average resampling—had higher vertical uncertainties.

The TanDEM-X and recently updated SRTM-C were both referenced to high-accuracy ICESat (Schutz et al., 2005; Zwally et al., 2009) measurements (collected between 2003–2009) during final block adjustments (Crippen et al., 2016; Rizzoli et al., 2017). While this removes the complete independence of these datasets, the relative sparsity of these points (170 m along track and up to 80 km across track) does not provide a continuous adjustment surface, but rather acts to improve local elevations and overall DEM quality with respect to remaining tilts (Rizzoli et al., 2017). Throughout the study *dh* refers to the TanDEM-X−SRTM-C 15 year differences (including both real change and vertical uncertainties).

# SRTM-C Correction Steps

Our correction chain was applied using the previous SRTM-C output at each stage as input in the following step. All steps were carried out on a $1^\circ \times 1^\circ$ tile-by-tile basis (unprojected WGS84 vertical and horizontal datums), however, merging tiles and then processing produced identical results. We also found comparable results using Universal Transverse Mercator (UTM) equal area projected tiles. The correction steps served to correct SRTM-C orbital biases and did not attempt to correct for terrain characteristics. We assumed that actual vertical change in our study area represented an extremely small fraction of pixels in the ~13 million pixel *dh* raster for each tile. This ensures that the corrections only rectified SRTM-C biases on stable terrain and were not influenced by smaller areas of true vertical land-level changes. Comparison of correction steps was done using normalized percentage difference histograms and quantile-quantile (QQ) plots.

## Co-registration

We corrected for sub-pixel offsets known to affect DEM comparisons (Van Niel et al., 2008; Berthier et al., 2007) using the universal co-registration of Nuth and Kääb (2011). This rigid translation is based on a cosine function fit to the relationship between terrain aspect and *dh* normalized by terrain slope:

$$\frac{dh}{\tan(\alpha)} = a \cdot \cos(b - \psi) + c \tag{3.1}$$

where $\alpha$ is slope, $\psi$ is aspect, and the variables $a$, $b$, and $c$ are the magnitude, direction, and mean bias, respectively. The shifts were applied to the SRTM-C by bilinear resampling with the $dx = a \cdot \cos(b)$ and $dy = a \cdot \sin(b)$ vectors used to weight the neighboring cells, and the mean shift $dz = c \cdot \tan(\bar{\alpha})$ added at the end.

We fit equation 3.1 to only slopes > 5° and, if necessary based on goodness of fit parameters, continued iteration of the fitting, shift vector solving, and interpolation until the magnitude of the shift vector ($a$) was < 0.5 m or the reduction in normalized median absolute difference (NMAD; Höhle and Höhle (2009)) on stable terrain was < 5% (Nuth and Kääb, 2011).

Our co-registration did not correct for slope and curvature using polynomial fitting (e.g., Kääb, 2005; Gardelle et al., 2012) as this introduces empirical models and additional uncertainties. We did not observe a linear positive or negative trend between slope and *dh* (Fig. A7). Curvature versus *dh* demonstrates the difference in actual resolution of raw sensor data between the SRTM-C and TanDEM-X (Fig. A10), however, correction of this intrinsic measurement limit introduces artificial elevations and are thus inappropriate for *dh* mapping between DEMs from different data sources and time-steps (see Appendix A).

Iterative shifting and bilinear resampling of one DEM to another by decimeter steps had the same effect on rectifying aspect biases (same shift vectors leading to minimization of bias) as the empirical fitting of the cosine relationship and calculation of shift vectors. This indicates the robust nature of the method of Nuth and Kääb (2011), assuming a sufficient distribution of high-slope, multi-aspect-facing topography is available for cosine fitting. The minimization of the sum of errors and cross-correlation methods (e.g., Kääb, 2005) were unsuccessful at removing shifts in our study region.

## Destriping

For removal of long- and short-wavelength striping patterns in the SRTM-C, we followed previous work using frequency analysis techniques to identify striping artifacts (e.g., Arrell et al., 2008) and noise (e.g., Purinton and Bookhagen, 2017) in DEMs. We took particular inspiration from Yamazaki et al. (2017) and used fast fourier transforms (FFTs) to filter the *dh*. In a first step, we removed all pixels identified as having low coherence in the TanDEM-X WAM. This filtered large water bodies and other areas that may show artifact noise affecting FFT analysis. Following this, any void pixels (including the low-coherence areas) were set to $dh = 0$ and an FFT was run. The power spectral density (PSD) was calculated as the magnitude of the FFT squared and a mean 5×5 filter was passed over it. The ratio of original and smoothed PSD was then taken to identify regions of the spectrum with high outliers (high ratio) representing cyclic, tile-spanning stripe bias. We used the $97.5^{th}$ percentile of the ratio as the cutoff value. The remaining top 2.5% high- and low-frequency outliers received an inverse FFT, which produced a map of the long- and short-wavelength stripes. These stripes were then removed from the SRTM-C and the process was repeated iteratively until the improvement in root mean squared error (RMSE) was < 5%.

We refer to the above parameters as non-aggressive destriping, since we are just "shaving off" the top of the distribution. In aggressive tests, we experimented with lower percentile cutoff values (e.g., $95^{th}$) and lower tolerance for RMSE convergence (e.g., < 2% improvement). While these more aggressive destriping schemes did successfully eliminate the SRTM-C orbital biases, we also found that the true topography was often filtered following the > 5

iterations needed to meet the RMSE convergence requirements (Fig. A11). Therefore, we chose to use the non-aggressive cutoffs and ran additional blocked shifting discussed in the following section.

### Blocked Shifting

Patchy positive and negative regions in the co-registered, destriped *dh* map were solved by breaking the $1^\circ \times 1^\circ$ tile into square blocks and shifting each block by the median value. These areas likely correspond to remaining orbital biases that were not removed in our non-aggressive destriping technique. There may be local correspondence between these patches and atmospheric water vapor conditions at the time of SRTM-C collection in February 2000, however, such data at the sub-kilometer scale necessary for analysis is unavailable. Furthermore, local adjustment of the SRTM-C and TanDEM-X to ICESat measurements could contribute to these shifts, though the contribution is difficult to quantify.

We began by masking the low-coherence pixels (again from the WAM) since these would disproportionately contribute to local median shifts. Using a variety of block sizes with edge lengths ranging from 1.35–7.2 km, we found the median *dh* and median slope in each block. We used the median slope to normalize the median *dh* values, since we expect areas of higher slope to have greater uncertainties and biases (Fig. A7) unrelated to SRTM-C orbital biases. Furthermore, we allowed a maximum shift per block of $\pm 1$ m, thus ensuring that this step did not cause unreasonably large shifts due to outliers contained in a given block.

## Differencing for Change Detection

Following orbital SRTM-C bias corrections, it is possible to merge corrected tiles and create maps of *dh* to measure areas of actual change. Previous change mapping over gravel-bed channels has relied on level of detection cutoffs and probabilistic thresholding (e.g., Lane et al., 2003; Wheaton et al., 2010). These studies have, however, been developed for meter to sub-meter photogrammetric or lidar data. Here we use a hybrid approach of statistical outlier detection on the entire distribution of pixels followed by a level of detection cutoff for remaining pixels well within the bounds for expected noise between the datasets. Remaining uncertainties are primarily caused by speckle noise and terrain characteristics, with the biggest impact from slope. The following sections provide a detailed description of the change detection method for channels and hillslopes.

### Channels

We know from field observations that large braided gravel-bed channels in the study area (Fig. 3.1B) change rapidly with local incision and aggradation (natural and anthropogenic in the form of gravel mining) on the order of meters during the past decade. Outlines of the bank-to-bank active width of the primary channel branch were digitized from open-source satellite imagery from Bing$^{TM}$ and GoogleEarth$^{TM}$. We buffered the resulting channels by $-60$ m (upper limit of gridded SRTM-C resolution). This means we only use the wide (> 120 m), non-vegetated channel reaches from Río Toro and Río Grande where there has been recent aggradation and incision.

Change mapping was done by separating the in-channel *dh* values into bins of contributing error factors (local relief and TanDEM-X individual scene consistency) and applying $5^{th}$ and $95^{th}$ percentile cutoffs to each bin, thus only taking the top (positive=aggradation) and bottom (negative=incision) 5% of outliers. We first used the TanDEM-X WAM to remove the untrustworthy *dh* pixels where coherence was lost three or more times (Wessel,

2016). Because gravel-bed channels represent a low-slope environment with no vegetation and we are only measuring wide valleys, we assumed that DEM error from SRTM-C and TanDEM-X were restricted to random speckle noise. Nonetheless, to account for steeper areas with potentially more error from phase unwrapping, we separated *dh* into relief bins using the pixels' 500-m radius relief values. We also separated *dh* by the TanDEM-X consistency (COM) and height error (HEM) masks (Fig. A2–A3). Taken together, *dh* pixels in high-relief, high-height error, and low-consistency bins required greater magnitudes to avoid noise cutoffs than vice versa. A minimum level of detection approach (Lane et al., 2003) was taken as the RMSE of the entire *dh* map on low-slope (similar to channel slope) areas. In a final step, all remaining in-channel *dh* values below this RMSE cutoff were removed as likely noise. Volume changes are calculated from the sum of pixel area (900 $m^2$) multiplied by vertical change, with uncertainties taken as the level of detection RMSE and propagated via equation (15) in Lane et al. (2003).

### Entire Landscape

When considering *dh* over the entire landscape, we include far more uncertainties related chiefly to steeper terrain. Thus, the error must be handled differently than for strictly low-slope pixels (in-channel). First, a corrected *dh* map for the entire study area was generated. Similar to channel mapping, low-coherence pixels were removed with the WAM and *dh* was separated into bins of slope, height error, and consistency to retrieve only the top and bottom 5% of outliers in each bin set. The level of detection cutoff was taken as the RMSE across the entire landscape, which was almost entirely stable terrain, and remaining *dh* values below this cutoff were eliminated.

At this stage, a great many lone and patchy *dh* values remained. Given this, it was not possible to automatically identify areas of change that were only a small number of pixels in size. Interested in large-scale changes, likely not associated with a single pixel, we sought connected pixels showing all up or all down vertical motion. To winnow the potential change pixels, we applied binary opening with a 1-pixel radius circular kernel, thus removing many unconnected outliers and small patches. Next, we took the summed *dh* of each separate patch. It was assumed that the majority of patches, and thus majority of summed values, were remaining noise in the difference map, whereas signal should be spatially coherent and largely positive or negative. Therefore, by applying a standard deviation cutoff over summed patches (here we used 1-$\sigma$, though this can be easily set for testing), we removed a vast majority of remaining pixels, and only kept the largest outliers. This limited the method to only assessing the largest coherent vertical changes in the landscape, but eliminated the possibility of mis-identifying change that was in fact noise. These remaining patches can be explored in map-view and compared with satellite or historical imagery for further confirmation and analysis.

# Results

## Correction Steps

Co-registration of SRTM-C to TanDEM-X revealed X-Y shifts of ~1/10 of a pixel (~3.7 m). Although minor Z shifts (~1 m) were also determined and corrected during co-registration, these were not unique across entire tiles, but rather related to long-wavelength SRTM-C biases. The cosine fitting to *dh* normalized by terrain slope can be seen in Figure 3.2, whereas, in map-view the change is more subtle and difficult to discern.

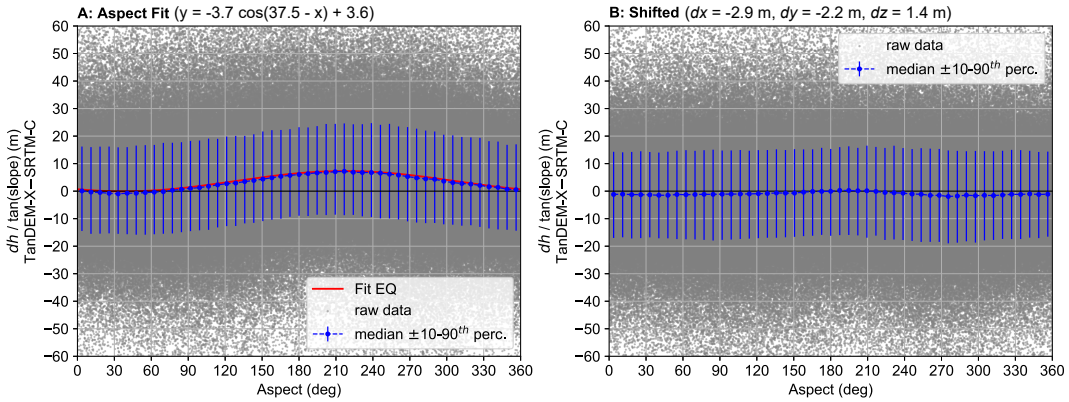


**Figure 3.2** – Relationship of $dh$ (normalized by tangent of slope) to aspect (A) before and (B) after co-registration and bilinear resampling of SRTM-C. We fit to equation 3.1 on all raw data. Note the close match between equation fit and median values. The cosine relationship in (A) is caused by overestimation of the SRTM-C on NE facing aspects (peaking at ~60°) and underestimation on SW facing aspects (peaking at ~220°). The resulting $(dx, dy)$ shift vector is directed SW.

In Figure 3.3, we demonstrate one iteration of destriping for a single SRTM-C tile (S 24°, W 66°). It is apparent in the co-registered $dh$ map that a number of long- and short-wavelength shifts are affecting the tile. Using our FFT, statistical cutoffs, inverse transform, and stripe removal, the resulting $dh$ map has a much more uniform appearance and the median and RMSE are both reduced. This process was typically repeated 2–4 times per tile, until the RMSE began to converge. While topographic uncertainties remain in steep and high-relief regions, the overprinting biases are reduced.

Since we do not use an aggressive FFT filtering scheme, a number of patchy outliers remain. We attempted to correct these regions using blocked shifting (Fig. 3.4), shown in this case over three tiles covering the foreland and Altiplano-Puna Plateau Region (S 24–26°, W 66°). After testing multiple block sizes, we preferred blocks with edge length of 3.6 km, since these provide a small enough area to correct highly localized inconsistencies, while also being far greater in size than the largest vertical changes we would expect in the landscape.

## Comparison of Correction Steps

Since stacked histograms are difficult to interpret and larger magnitude outliers are fewer in number and thus obscured, we plotted the normalized bin percentage difference of $dh$ in each step of correction (Fig. 3.5). Co-registration mostly caused a mean shift in the distribution. Moving to destriping, the number of pixels at high outlier values went down significantly (> 20% drop in ±15–20 m bins) and there was some (~10%) increase in bins ±5 m, whereas the number of values close to zero $dh$ decreased. This represents an overall re-distribution of error from the SRTM-C orbital biased patterns (Fig. 3.3) to a more uniform spatial pattern (Fig. 3.4). The final blocked shifting caused very little overall change in the distribution, which was mostly in the form of another mean shift (this time directed the other way from co-registration). These effects can also be seen in a QQ plot of each subsequent correction step (Fig. 3.6), where co-registration caused a mean shift and some outlier reduction, de-striping had a large effect on narrowing the distribution at the tails, and blocked shifting again had a minimal effect on nar-

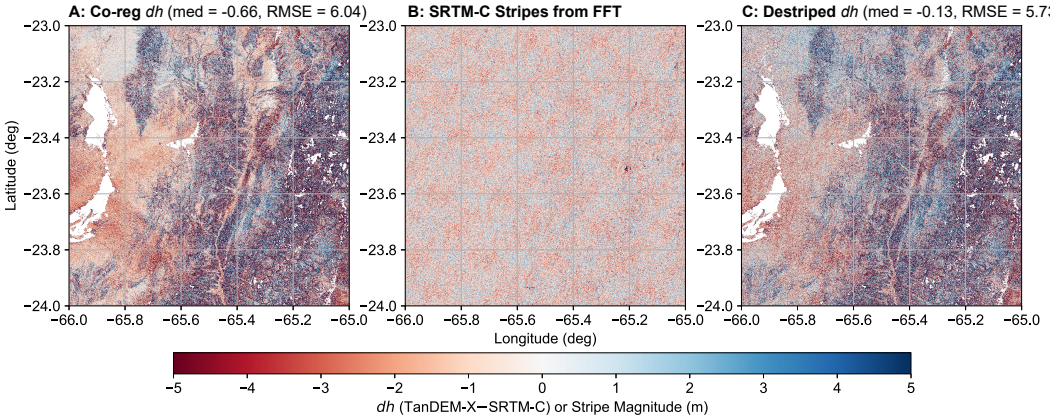


**Figure 3.3** – One iteration of FFT destriping from one tile (S 24°, W 66°). Both median and RMSE improve from (A) the co-registered map to (C) the destriped map. Stripes removed by FFT are shown in (B). Note that (C) is not the final corrected map as iteration was run twice more before RMSE began to converge at 5% tolerance level. Voids (white space) are untrustworthy pixels removed by TanDEM-X WAM cutoff prior to destriping.

rowing the distribution at the most extreme outliers. In all cases, the median value (0.5 quantile) moved closer to zero. Overall, these plots indicate the importance of SRTM-C correction and of the destriping step in particular prior to using TanDEM-X−SRTM-C *dh* maps for change mapping.

# Areas of Change

As discussed in the methods, we separated potential change identification and measurement from corrected (co-registered, destriped, block shifted) *dh* maps between the in-channel pixels and the entire landscape.

## Channels

Binning corrected in-channel *dh* and cutting off any remaining outliers within the low-slope RMSE of ~3 m reduced the data density significantly by cutting out any pixels within expected noise. The potential signal pixels were then plotted atop longitudinal profiles from the Río Toro and Río Grande (Fig. 3.7). The point clouds of *dh* values were colored with a Gaussian kernel density estimate (KDE) to demonstrate the denser (warmer colors) versus sparser (cooler colors) zones of measurement. The density is displayed as percentiles of the full distribution of the 2D KDE of *dh* from both channels. Turning to map-view, we can observe the location of these pixels in the channel and their relation to local characteristics, upstream factors, and anthropogenic tampering (Fig. 3.8).

## Entire Landscape

To be mapped as true vertical change, an area in the greater landscape must be significantly large and coherently positive or negative since many of the pure noise patches are > 10 pixels in size (> 0.01 $km^2$). Furthermore, the individual pixels must show significant height changes above the overall RMSE of ~6 m and outlier cutoffs in each bin, which in steeper bins may be > 10 m. Examining results in map-view (Fig. 3.9) allows assessment of the potential true signal versus noise. At this stage it is necessary to include auxiliary data from field knowledge or remote sources like aerial or satellite imagery (e.g., GoogleEarth$^{TM}$). Our method was able to identify one major landslide in

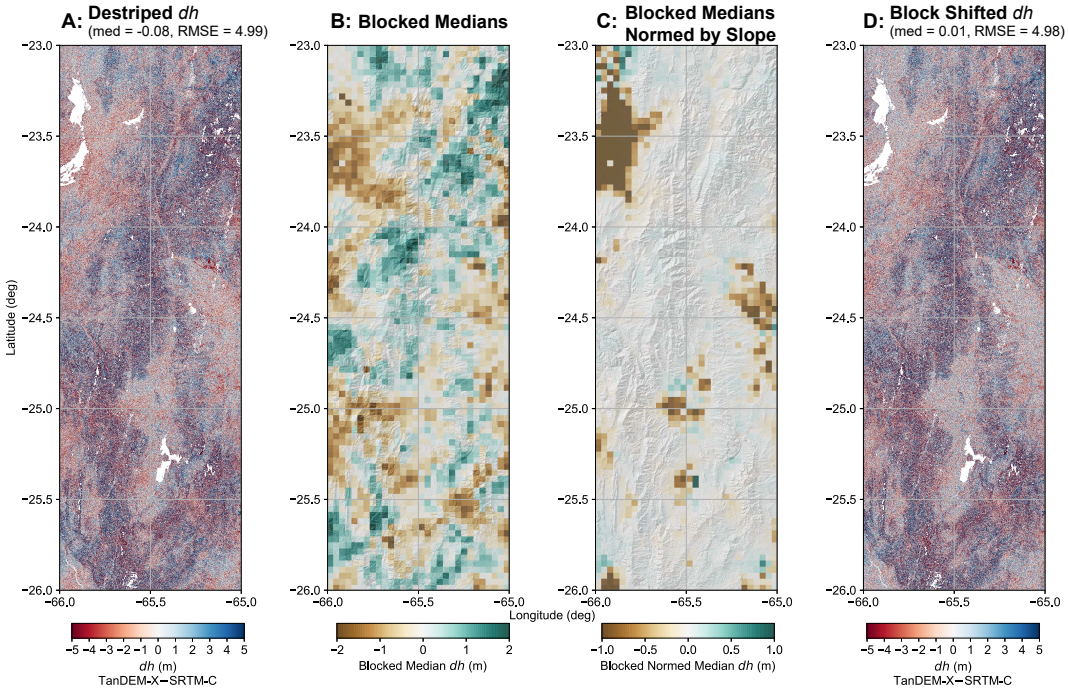


**Figure 3.4** – Blocked shifting on three destriped and merged tiles (S 24–26°, W 66°). Blocks are 3.6 km in height and width. The (A) destriped median and RMSE both improve slightly in (D) the final shifted *dh* map. Note that the original blocked medians (B) show a slight pattern resembling the long-wavelength stripe bias from SRTM-C. In (C) we have normalized the median shifts by the median slope values, so as not to over-correct the steeper regions with higher uncertainties. The color scheme is changed for (B) and (C), and the scale of (C) is half the width of (B) since it only extends to the maximum allowable shift of ±1 m. Scales and color scheme in (A) and (D) are identical. Voids (white space) are untrustworthy pixels removed by TanDEM-X WAM cutoff prior to median calculation.

the study area (Fig. 3.9D), however, most other measurements are remaining large artifacts attributable to both the SRTM-C and TanDEM-X. Low-coherence zones that may represent change between TerraSAR-X / TanDEM-X contributing scene collection (Fig. 3.9B–C) are necessarily removed in the WAM cutoff prior to binning.

# Discussion

## Necessity of Correction Steps

The original SRTM-C is plagued by numerous terrain and sensor specific errors and biases (e.g., Carabajal and Harding, 2006; Gorokhovich and Voustianiouk, 2006; Van Niel et al., 2008; Gallant and Read, 2009; Yamazaki et al., 2017). Despite re-processing of the original data in the new NASADEM product, many of these errors remain (Crippen et al., 2016). On the other hand, the newer TanDEM-X apparently has far fewer biases related to satellite geometry, and most error is restricted to terrain characteristics like slope and vegetation, though results are still nascent (e.g., Baade and Schmullius, 2016; Purinton and Bookhagen, 2017; Wessel et al., 2018). Our correction steps do not seek to eliminate bias related to terrain characteristics at the scale of a few hundred meters, but rather

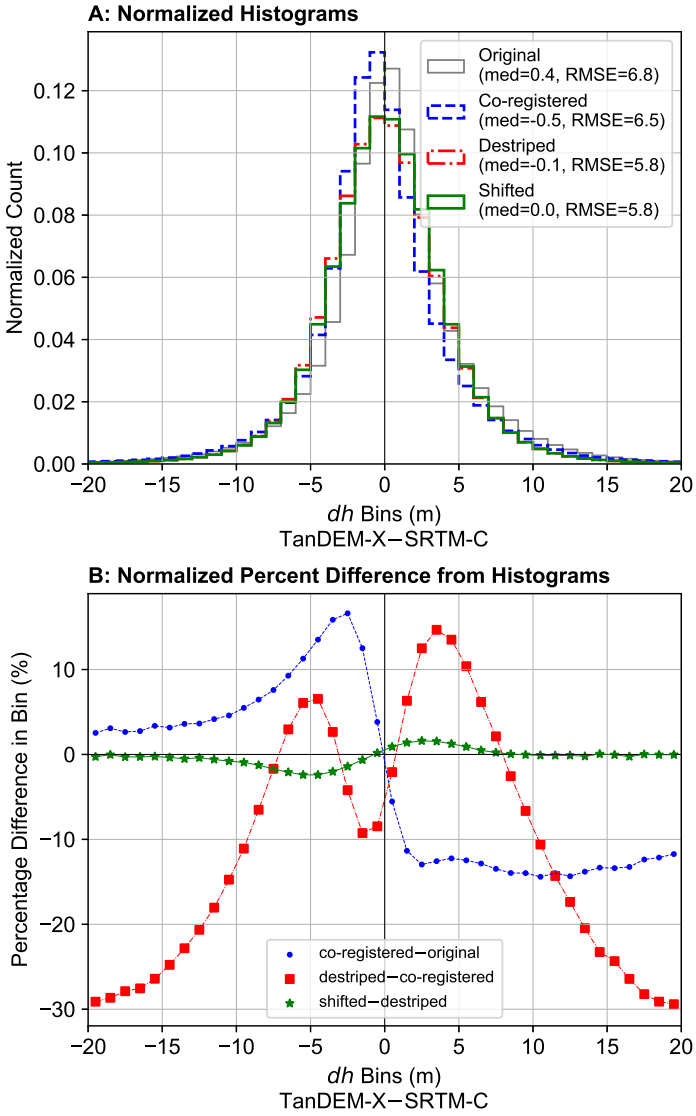


**Figure 3.5** – Characteristic (A) stacked histograms and (B) normalized percentage bin difference from three tiles merged and processed (S 24–26°, W 66°). Though it is difficult to interpret the histograms, plotting their difference (normalized by bin count) as percentage change between successive steps demonstrates the shifting of the median to near-zero and the reduction in outliers.

to correct large scale biases related to primarily the SRTM-C at scales of several hundred meters to kilometers. Perhaps this reduction in bias is most obvious in map-view of the subsequent *dh* patterns between processing steps (Fig. 3.3A to Fig. 3.4A to Fig. 3.4D), but we also show statistically that these steps lead to a narrowing of the distribution and centering of the differences on zero-median (Fig. 3.5–3.6). We assume that the vast majority of the pixels (outside of the cryosphere) should be unchanged over 15 years, and thus median shifts between the datasets at large scales are biases in need of correction.

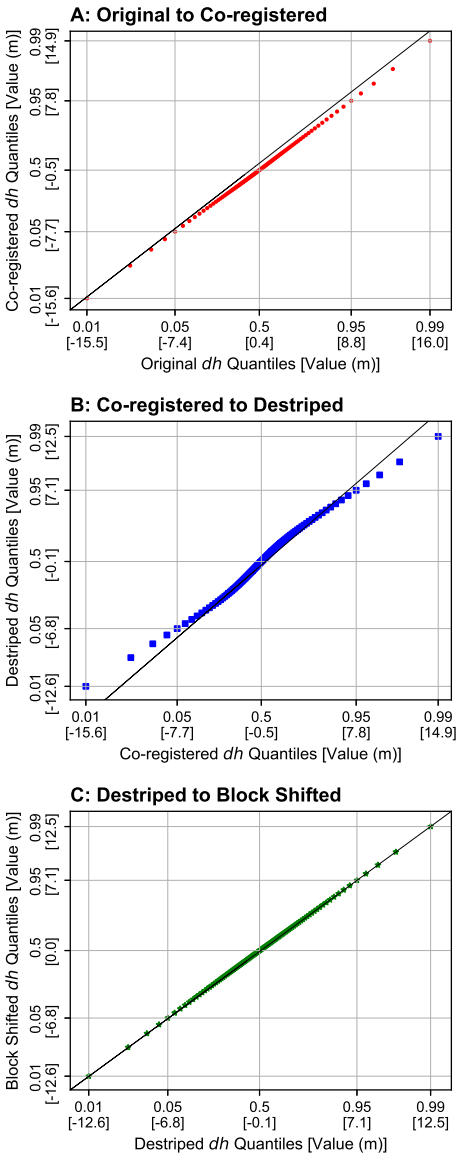


**Figure 3.6** – Quantile-quantile (QQ) plots showing difference between each successive correction step from three tiles merged and processed (S 24–26°, W 66°). (A) original to co-registered, (B) co-registered to destriped, and (C) destriped to block shifted. We note that co-registration and destriping have the greatest effect on zero-median-shifting and narrowing the outliers. The quantiles (0.01, 0.05, 0.5, 0.95, and 0.99) and their respective values are indicated on each axis to highlight this effect.

Co-registration indicates NE facing aspects are overestimated by the SRTM-C causing a negative excursion in the cosine fit, whereas SW facing aspects are underestimated and thus the *dh* compared to TanDEM-X is positive. This error mostly affects higher slopes (Nuth and Kääb, 2011), which is the reason for normalization of *dh* by the tangent of slope. The directions of bias correspond to the look direction orthogonal to the SRTM-C descending path and parallel to the ascending path. This indicates that the source of this bias is the SRTM-C, as reported by previous authors (Bourgine and Baghdadi, 2005; Gorokhovich and Voustianiouk, 2006; Shortridge and Messina, 2011), and

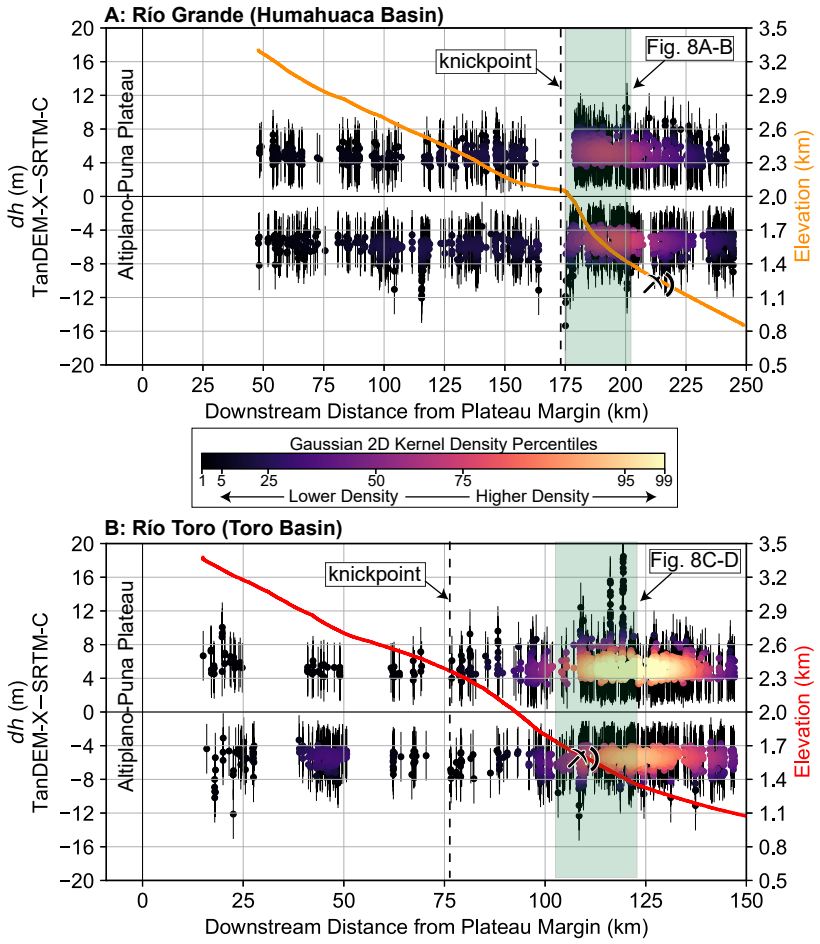


**Figure 3.7** – Longitudinal profiles of (A) Río Grande and (B) Río Toro overlain with point cloud of potential *dh* signal (pixels outside of the range of expected noise). Error bars are RMSE from low-slope (< 5°) terrain outside of the channel area. Each *dh* point cloud is colored by probability density from a Gaussian 2D KDE to show the denser (warmer) versus sparser (cooler) reaches. The KDE is scaled over all measurements from both channels and relative percentiles of the full distribution are used to highlight denser zones, particularly in (B) Río Toro. Note the $x$-axis range is 100 km greater for the longer Río Grande, despite the same axis scaling. Color scheme for elevation profiles on right axes match map-view color of each channel in Figure 3.1B. The knickpoint in Río Grande is caused by the large Del Medio fan (Savi et al., 2016), whereas the origin in Río Toro is tectonic, caused by the Gólgota Fault (Marrett et al., 1994; Hilley and Strecker, 2005). In both cases, the majority of the *dh* signal appears downstream of the knickpoint. Map-view of green highlighted regions is shown in Figure 3.8.

not TanDEM-X. A shift—accompanied by bilinear resampling—of just ~3.7 m (magnitude $a$ of equation 3.1 fit) to the SW rectifies this aspect bias.

As opposed to Yamazaki et al. (2017), we do not set a user defined ratio for FFT destriping, but rather use statistical "shaving off" of only the outlier stripe noise until the data converge. This conservative approach retains the true topographic signal at the expense of remaining stripe noise. In the case of more aggressive FFT filtering, using lower percentiles for the ratio cutoff and more strict RMSE convergence requirements, the actual topography began to

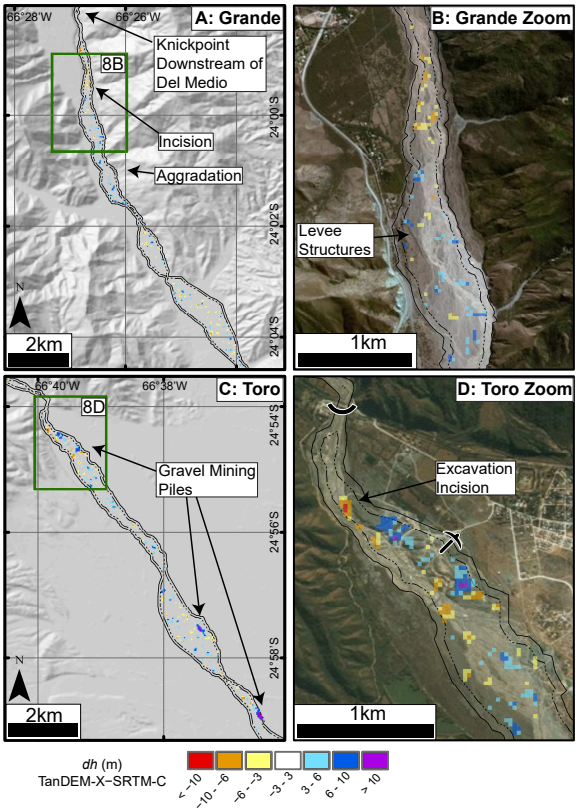


**Figure 3.8** – Map-views of the in-channel *dh* measurements for Río Grande (A) and Río Toro (C) highlighted in the longitudinal profiles in Figure 3.7. For location of each map refer to Figure 3.1C. More details are shown in zoom-ins of the in-channel *dh* measurements in (B) and (D). The solid outline is the digitized bank-to-bank channel and the stippled line is the −60 m buffer area of measurement. We note large areas of incision related to the steep and narrow channel downstream of the Del Medio fan and knickpoint in Río Grande (A), immediately followed by a zone of aggradation with levee structures to direct gravels (B). For Río Toro (C) we highlight the anthropogenic influence of gravel mining generating large piles and also causing incision due to local excavation (D).

filter out of the *dh* maps (Fig. A11), which, as stated, is not the aim of our orbital bias correction steps and would lead to the inclusion of artificial (i.e., FFT generated) *dh* measurements.

Remaining stripe noise is apparent in Figure 3.4B, where the blocked medians resemble the original long-wavelength stripe pattern, though discontinuous. Despite the appearance in some areas of more negative values in the western parts of tiles (higher elevation, Altiplano-Puna Plateau), we do not find any clear relation between block medians and elevation at any block size or in any tile (see Appendix A). Block shifting removes the remaining noise, but again we avoid correcting for strongly overprinting topographic biases related to slope by normalizing the block median *dh* by median slope. Overall, these steps provide a more trustworthy *dh* map, while respecting the inherent and difficult to account for biases in radar derived spaceborne DEMs.

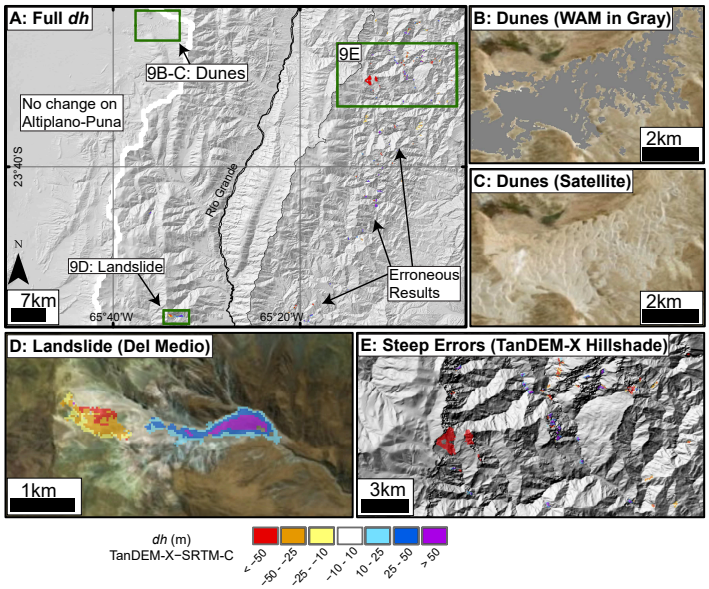


**Figure 3.9** – (A) Map-view of landscape-wide *dh* identification. For location refer to Figure 3.1C. Our method returns little change on the low-erosion Altiplano-Puna. The dunes (B–C) are not identified since they are masked out using the TanDEM-X auxiliary WAM as low-coherence zones. This indicates their rapid displacement between TerraSAR-X / TanDEM-X scene collection. Our method is able to identify one major landslide (D) in the Del Medio catchment (Savi et al., 2016), however, there are many erroneous results in steep and vegetated zones to the east, shown in (E) over the TanDEM-X hillshade.

# Potential Change Mapping

For lower slope regions (i.e., channels), the potential for change mapping is greater than in steeper areas. This is caused by the better agreement and lower vertical uncertainty of the two datasets in flatter, vegetation free areas. In both channels, the largest density of measurements is found below the respective knickpoints. This corresponds to an order of magnitude increase in the 2D KDE shown by the warm colored patches in Figure 3.7. In terms of the actual number of measurements (number of *dh* pixels) per binned channel reach, Figure A13 demonstrates this approximately five to ten fold increase in the downstream reaches with a simple histogram. This result partially has to do with a narrower channel and thus less measurements available above the knickpoints (hence the numerous gaps in measurement in the upstream reaches), however, these results also appear to indicate that the most geomorphic work is happening downstream of the oversteepening point. This also coincides with a transition to a wetter environment in both cases.

The Río Toro has a particularly dense zone of measurements at the mountain front where naturally high rates of aggradation are enhanced by human gravel excavation and piling. On the other hand, in the Río Grande the downstream measurements are spread over a greater channel reach and thus appear less dense in the 2D KDE (the measured Río Grande is ~100 km greater in length than the Río Toro). Downstream of the knickpoint, Río Toro is in a net aggradation state with a corrected *dh* volume of $0.81 \pm 0.15 \times 10^6$ m$^3$, whereas, for Río Grande the net state is incision with a volume of $-0.69 \pm 0.15 \times 10^6$ m$^3$. In comparison, the pre-correction volume in each case is

$-1.18\pm0.12\times10^6$ m$^3$ and $2.80\pm0.11\times10^6$ m$^3$ for Río Toro and Río Grande, respectively, thus indicating a flip in sign and reduction of magnitude following careful corrections applied prior to differencing.

Locally, the aggrading and incising patches may be related to braided channel avulsion and subsequent rapid incision into the unconsolidated bed material during frequent high-discharge events brought by convective rainfall in the summer monsoon (Castino et al., 2016a,b, 2017). In map-view (Fig. 3.8), we see that these automated measurements can be correlated with additional sources. For Río Grande, the steep knickpoint at the Del Medio fan (Savi et al., 2016; Schildgen et al., 2016) causes a major zone of incision immediately followed by aggradation where the material is deposited. Fieldwork has indicated that some of this incision is man-made, caused by attempted removal of aggrading material coming from the productive (e.g., debris flows cf. Savi et al. (2016)) Del Medio catchment. Levee structures (Fig. 3.8B) are a testament to this tendency towards aggradation downstream of this extremely erosive fan. The cause of aggradation in the Río Toro is clearly enormous gravel piles being created just at and downstream of the mountain front. The volume of the large gravel pile indicated in Figure 3.8D directly at the mountain front in Río Toro is $0.78\pm0.06\times10^6$ m$^3$, with this growth between SRTM-C and TanDEM-X observed during field work over the past decade and from GoogleEarth$^{TM}$ historical imagery back to 2003. This is coupled with incision in the active channel upstream of the piles where gravel is being removed to prevent widespread aggradation.

In terms of rates of change, our minimum measurable *dh* of $\pm3$ m corresponds to a rate of $\pm0.2$ m/yr, given the conservative 15 year time difference between DEMs. This rate represents an average for the entire measurement period and assumes constant geomorphic change, whereas the true rates are more stochastic, following rainfall and anthropogenic activity variation. The area of greatest point density in the longitudinal profiles in Figure 3.7 is centered at $\pm5$ m, corresponding to a rate of $\pm0.33$ m/yr, with maximum rates of incision and aggradation, occurring at anthropogenic gravel piles and excavation sites, in excess of $\pm0.5$ m/yr. Human tampering is known to cause significant excursions from natural river dynamics (Kondolf, 1997; Grant, 2012), and we have shown that signals of excavation and piling are highlighted as above-the-noise outliers. Previous studies have demonstrated similar rates over longer time-scales (tens to hundreds of years) using more sparse measurements (e.g., Rinaldi and Simon, 1998; Rovira et al., 2005; Walter and Merritts, 2008; Comiti et al., 2011) and at shorter time-scales (< 5 years) from meter-scale lidar data (Lane et al., 2003; Wheaton et al., 2010). The identification and quantification of incision and aggradation has important implications for infrastructure and agriculture given that 60% of global sediment delivery to coasts originates in high mountain regions (Syvitski et al., 2005).

Mapping *dh* signals across the entire landscape presents a greater challenge given the higher uncertainties on steeper more complex topography. Nevertheless, using the binning method, binary operations, and outlier selection removes a large portion of the noise from the corrected data. Our method displays very little change on the low-relief, low-slope Altiplano-Puna besides some salt flat areas that were not removed by the coherence masking from the TanDEM-X WAM. Remaining noise mapped as potential change is clear at the mountain front where steep slopes and heavy vegetation causes complication of accurate radar measurement. In many locations these erroneous patches correspond with low-amplitude or low-coherence zones also identified in the WAM. We were able to automatically map one landslide, previously reported on by Savi et al. (2016), in the Del Medio sub-catchment

of the Humahuaca Basin using this method. This material likely contributes to the aggradation we see occurring downstream of the fan in the longitudinal profile (Fig. 3.7A) and in map-view (Fig. 3.8A). The calculated detachment and deposit volumes from this massive earth movement are $-10.5\pm0.12\times10^6$ m$^3$ and $16\pm0.15\times10^6$ m$^3$, respectively, with vertical land-level changes greater than $\pm$50 m associated with the break-off and lobe (Fig. 3.9D). These magnitudes of change show little difference in the pre- and post-corrected mapping, indicating (a) this is a localized region of good agreement between SRTM-C and TanDEM-X and (b) this large landslide can be identified in uncorrected difference maps.

The area of sand dunes, clearly visible as a low-coherence region from the TanDEM-X WAM in Figure 3.1C and Figure 3.9B–C, is not mapped as potential change since the coherence masking prior to binning eliminates this area from consideration. Examination of *dh* in this region is very noisy since the TanDEM-X contains measurements spanning 5 years, thus causing completely different height inputs for the same pixel in many scenes. This indicates the potential of the WAM alone for mapping change on shorter time-scales outside of very steep areas.

## Caveats of Data and Method

Spaceborne DEMs present significant challenges for accurate height measurements, though until lidar or sub-meter satellite data becomes more widespread and cheaper (Passalacqua et al., 2015), it is the only option in many study areas. On the other hand, unmanned aerial vehicles and point clouds generated using structure from motion technology could already provide a viable alternative (Javernick et al., 2014; Cook, 2017), but applying these methods at the scale of entire catchments or over tens-of-kilometers of river reaches is not feasible. Previously, *dh* measurement from space has been primarily focused on the cryosphere (e.g., Berthier et al., 2006; Nuth and Kääb, 2011; Neelmeijer et al., 2017) due to limitations in data accuracy. Certainly radar data are more adequate than optical data (e.g., Fisher et al., 2013; Purinton and Bookhagen, 2017) for the case of unconsolidated sediment, particularly since different penetration depths do not affect measurement (Rignot et al., 2001; Rossi et al., 2016), assuming limited vegetation.

Here we have demonstrated the potential of new high-accuracy datasets such as TanDEM-X to correct outstanding biases in the SRTM-C and potentially contribute to land-level change mapping and measurement over previously unattainable scales. Given remaining noise in the datasets, change mapping is limited to large areas of coherent change (e.g., massive landslides) or specific low-slope areas of interest such as wide gravel-bed rivers. In any case, field data (e.g., repeat total station or GPS surveys), field knowledge (e.g., via observations of incising reaches or roads damaged by aggrading channels), and/or auxiliary data (e.g., GoogleEarth$^{TM}$ historical imagery change mapping) are necessary for accurate assessment of the location of true change signals versus noise. Further, the magnitude of change must be significantly above the expected uncertainty between DEMs, which in the case of SRTM-C and TanDEM-X is as low as ~3 m on flat, sparsely vegetated terrain, and increasing with slope and topographic complexity. We posit that these correction steps may also be applied to cryospheric studies, however, radar penetration would need to be carefully considered first as this may exceed *dh* signals.

# Conclusions

In this study we have presented a novel use of two near-global spaceborne DEMs (SRTM-C and TanDEM-X) separated by ~15 years to measure land-level changes in the south-central Andes in northwestern Argentina. Previous measurement of land-level changes at the scale of entire mountain belts has been restricted to the cryosphere, where the signal of snow and ice change outweighs the noise associated with DEMs used for differencing (typically ASTER or single TerraSAR-X / TanDEM-X CoSSC DEMs). On the other hand, studies outside of the cryosphere have relied on high-accuracy meter to sub-meter data at much smaller scales to measure height changes in rivers and hillslopes. Using the TanDEM-X DEM as a control surface, we corrected long-standing SRTM-C errors related to orbital biases. We then successfully differenced the two datasets to identify and quantify land-level changes outside of expected noise caused by radar DEM speckle and other terrain dependent errors, increasing with steep and complex topography. Noise from imperfect datasets continues to hinder signal detection in low magnitude geomorphic change detection, however, this study continues to push the envelope of the potential for change mapping using the data currently available to many scientists.

Our method is useful for the case of large gravel-bed rivers where the width far exceeds SRTM-C 1 arcsec resolution considerations. In such flat, vegetation free environments it is useful to analyze the river alone and not include additional uncertainties brought by increasing slopes and dense vegetation. For these steeper regions, the use of greater outlier cutoffs and the necessity for large and coherent patches of land-level change, both to remove the majority of noise, limits the method to only very large earth movements. In either case, only signals outside of expected noise can be confidently identified, which in the case of gravel-bed rivers typically fall in the realm of human tampering. From the TanDEM-X auxiliary data alone it is also possible to identify regions that changed during TanDEM-X collection (2010–2015) using the water indication mask, however, this does not provide quantifiable change.

Overall, the use of relatively coarse (1 arcsec) spaceborne DEMs to derive land-level changes benefit from higher accuracy radar-derived data, whereas the use of optical data is limited to sub-meter resolution satellites. The application of this method to other regions around the world could indicate previously unmapped vertical changes. In the future, both the SRTM-C and TanDEM-X will continue to be used as snapshots of the earth's surface separated by over a decade, and thus useful for differencing against newer datasets yet to be developed to continue measuring vertical change outside of the cryosphere.

# Code and Data Availability

Python codes for co-registration, FFT destriping, blocked shifting, and potential change mapping are available on GitHub at https://github.com/UP-RS-ESP/TanDEM-SRTM-dh.git. The SRTM-C updated NASADEM tiles can be found at: https://e4ftl01.cr.usgs.gov/provisional/MEaSUREs/NASADEM/. TanDEM-X data is only available from the DLR through scientific proposals.

# Acknowledgements

The authors thank the DLR for TanDEM-X DEMs received through DEM_CALVAL1028 to B. Purinton and DEM_GEOL1762 to Stephanie Olen. Additional funding was sourced from DFG Graduate School StRATEGy (IGK2018) and NEXUS funded through the MWFK Brandenburg, Germany, both to B. Bookhagen.

# Supporting Information

See Appendix A.

Chapter 4

# Introducing *PebbleCounts*: A grain-sizing tool for photo surveys of dynamic gravel-bed rivers[†]


## Abstract

Grain-size distributions are a key geomorphic metric of gravel-bed rivers. Traditional measurement methods include manual counting or photo sieving, but these are achievable only at the 1–10 $m^2$ scale. With the advent of drones and increasingly high-resolution cameras, we can now generate orthoimagery over hectares at mm to cm resolution. These scales, along with the complexity of high-mountain rivers, necessitate different approaches for photo sieving. As opposed to other image segmentation methods that use a watershed approach, our open-source algorithm, *PebbleCounts*, relies on k-means clustering in the spatial and spectral domain and rapid manual selection of well-delineated grains. This improves grain-size estimates for complex river-bed imagery, without post processing. We also develop a fully automated method, *PebbleCountsAuto*, that relies on edge detection and filtering suspect grains, without the k-means clustering or manual selection steps. The algorithms are tested in controlled indoor conditions on three arrays of pebbles and then applied to 12 $\times$ 1 $m^2$ orthomosaic clips of high-energy mountain rivers collected with a camera-on-mast setup (akin to a low-flying drone). A 20-pixel b-axis length lower truncation is necessary for attaining accurate grain-size distributions. For the k-means *PebbleCounts* approach, average percentile bias and precision are 0.03 and 0.09 $\psi$, respectively, for ~1.16 mm/pixel images, and 0.07 and 0.05 $\psi$ for one 0.32 mm/pixel image. The automatic approach has higher bias and precision of 0.13 and 0.15 $\psi$, respectively, for ~1.16 mm/pixel images, but similar values of $-0.06$ and 0.05 $\psi$ for one 0.32 mm/pixel image. For the automatic approach, only at best 70% of the grains are correct identifications, and typically around 50%. *PebbleCounts* operates most effectively at the 1 $m^2$ patch scale, where it can be applied in ~5–10 minutes on many patches to acquire accurate grain-size data over 10–100 $m^2$ areas. These data can be used to validate *PebbleCountsAuto* applied at the scale of entire survey sites ($10^2$–$10^4$ $m^2$). We synthesize results and recommend best practices for image collection, orthomosaic generation, and grain-size measurement using both algorithms.



[†] *Purinton, B. and Bookhagen, B.: Introducing PebbleCounts: a grain-sizing tool for photo surveys of dynamic gravel-bed rivers, Earth Surf. Dynam., 7, 859–877, https://doi.org/10.5194/esurf-7-859-2019, 2019.*

# Introduction

Gravel-bed rivers transport water, nutrients, and sediment downstream, linking high mountains to populated forelands. The grain-size distributions — and associated percentile diameters, such as the $D_{50}$ and $D_{84}$ — in a river reach are fundamental geomorphic metrics of these systems (e.g., Shields, 1936; Parker et al., 1982; Church et al., 1998). They are used to characterize aquatic habitats (e.g., Kondolf and Wolman, 1993), assess the impacts of human infrastructure like dams (e.g., Kondolf, 1997; Grant, 2012), calibrate theoretical models of river transport and erosion (e.g., Sklar et al., 2006; Attal and Lavé, 2006; Attal et al., 2015; Dunne and Jerolmack, 2018), and explore natural phenomena such as downstream fining (e.g., Paola et al., 1992; Ferguson et al., 1996; Rice and Church, 1998; Gomez et al., 2001; Chatanantavet et al., 2010; Lamb and Venditti, 2016), which is essential for nutrient transport and ecological diversity.

Accurate grain-size measurement is elusive in nature given the heterogeneity of gravel-bed rivers, particularly in steep mountain catchments where the range of grain sizes is large. Traditionally, grain-size distributions have been gathered via physical clast measurement and counting along grids (Wolman, 1954), lines (Wohl et al., 1996), or in ~1 $m^2$ patches (Bunte and Abt, 2001), all truncated at some lower observable limit (e.g., Rice and Church, 1998). Not only are these techniques time consuming, prone to operator bias, and disruptive to the environment, but they also require large (hundreds of pebbles) sample sizes to accurately estimate the characteristic nature of the grains in each location (Wolcott and Church, 1991).

In light of this, measurement from photographs is an attractive option for increasing sample size and decreasing fieldwork, while covering larger areas. Increasingly affordable high-resolution — 12–24 megapixel — cameras, allows the collection of high-quality photo surveys via Structure from Motion with Multi-View Stereo (SfM-MVS) (Smith et al., 2016; Eltner et al., 2016) at scales of entire river cross sections or reaches with resolutions at, or exceeding, 1 cm/pixel (e.g., Woodget and Austrums, 2017). Even higher resolution (1 mm/pixel) river surveys can be accomplished with low-flying unmanned aerial vehicles (UAVs) (e.g., Carbonneau et al., 2018), pole-mounted cameras, or using handheld imagery.

We build on previous work and introduce the addition of color-space clustering techniques to present efficient new semi-automated (*PebbleCounts*) and fully automated (*PebbleCountsAuto*) algorithms for grain sizing from imagery in high-energy mountain rivers. Our algorithms are built on Python with a few popular libraries and are open source. The instructions and code can be accessed at: https://github.com/UP-RS-ESP/PebbleCounts (Purinton and Bookhagen, 2019a). In this study, we present previous work on grain-size measurement from rivers and our motivation for new developments. The processing chains of *PebbleCounts* and *PebbleCountsAuto* are then discussed. We test the algorithms in controlled conditions and then in a more challenging field setting in the northwestern Argentine Andes. The limits and caveats of the method are discussed using imagery of varying resolution, and suggestions for photo collection and processing are provided.

# Prior Studies

Modern digital grain sizing is divided into texture- and segmentation-based image-processing methods, as opposed to previous manual digitization (e.g., Kellerhals and Bray, 1971; Ibbeken and Schleyer, 1986). Many texture methods

rely on the relationship between grains and their shadowed interstices to derive size estimates over image windows. Examples include semivariance (Verdú et al., 2005; Carbonneau et al., 2003, 2004; Carbonneau, 2005), entropy or inertia calculated from gray level co-occurrence matrices (Haralick et al., 1973; Carbonneau et al., 2004; Carbonneau, 2005; Dugdale et al., 2010; de Haas et al., 2014; Woodget and Austrums, 2017; Woodget et al., 2018), and autocorrelation (Rubin, 2004; Warrick et al., 2009; Buscombe et al., 2010). These methods only provide one estimate of grain size (e.g., $D_{50}$), which often requires site-specific calibration.

Buscombe (2013) achieved full grain-size distribution measurements using wavelet decomposition, and published an open-source Python tool, *pyDGS*. This is a texture method that has been designed for the analysis of thin sections or beach sands and requires each grain to be fully resolvable and the distributions to be fairly homogeneous in size and shape. Additional texture methods rely on the 3D texture (or roughness) of point clouds to relate the variance of bed-scale topography to average grain size (Brasington et al., 2012; Rychkov et al., 2012; Westoby et al., 2015; Woodget and Austrums, 2017; Bertin and Friedrich, 2016), however, these techniques also requires site calibration and the relationships have been found to vary widely (Pearson et al., 2017).

In contrast to texture methods, the focus of segmentation is the full delineation and measurement of every visible grain. Segmentation is error prone in images that contain overlapping grains, a large range of grain sizes including sand patches, changes in landcover (e.g., vegetation), pebbles that are highly irregular in shape (non-ellipsoid), pebbles with intra-granular color variations or texture such as veins or fractures, and in which shadowing is irregular. Herein, we refer to these factors collectively as image complexity. Furthermore, segmentation-based methods also require high-spatial resolution point clouds or images that resolve the specific grain geometries. The benefits are that segmentation does not require any site calibration besides knowledge of the image scale and it provides a full grain-size distribution and all the commonly used percentiles ($D_{5,16,25,50,75,84,95}$). Published methods by Butler et al. (2001), Sime and Ferguson (2003), and Graham et al. (2005a,b) all rely on edge detection followed by watershed segmentation and ellipse fitting to each separate grain to get the long (a) and intermediate (b) axes. Detert and Weitbrecht (2012) added some sophistication to the algorithm of Graham et al. (2005a,b) and provide a free — though closed source — application called *Basegrain* for *Matlab*$^{TM}$, which has become a standard tool (e.g., Bertin and Friedrich, 2016; Bertin et al., 2017; Langhammer et al., 2017; Carbonneau et al., 2018).

# Motivation

Watershed segmentation is effective for interlocking, uniformly colored, oblate grains, however, energetic gravel-bed rivers in mountains often have more complex grain compositions with intra-granular variation, irregular shadowing, and a large range of sizes. The automated watershed methods proposed suffer from over-segmentation, grain misidentification, and the need for significant, time-consuming post-processing (e.g., in *Basegrain* with the split, merge, and delete tools) when applied to complex images. These issues limit their application to areas < 10 $m^2$.

Thus, we are motivated to develop a new semi-automated technique that uses k-means clustering of pixels and rapid manual selection of well-defined grains, herein referred to as the K-means with Manual Selection (KMS) or *PebbleCounts* approach, and a fully automated version that uses filtering of suspect grains, herein referred to as the Automatic with Image Filtering (AIF) or *PebbleCountsAuto* approach (Fig. 4.1). By avoiding over-segmentation and misidentification, we are able to select fewer grains per image, but be sure that those selected are correctly

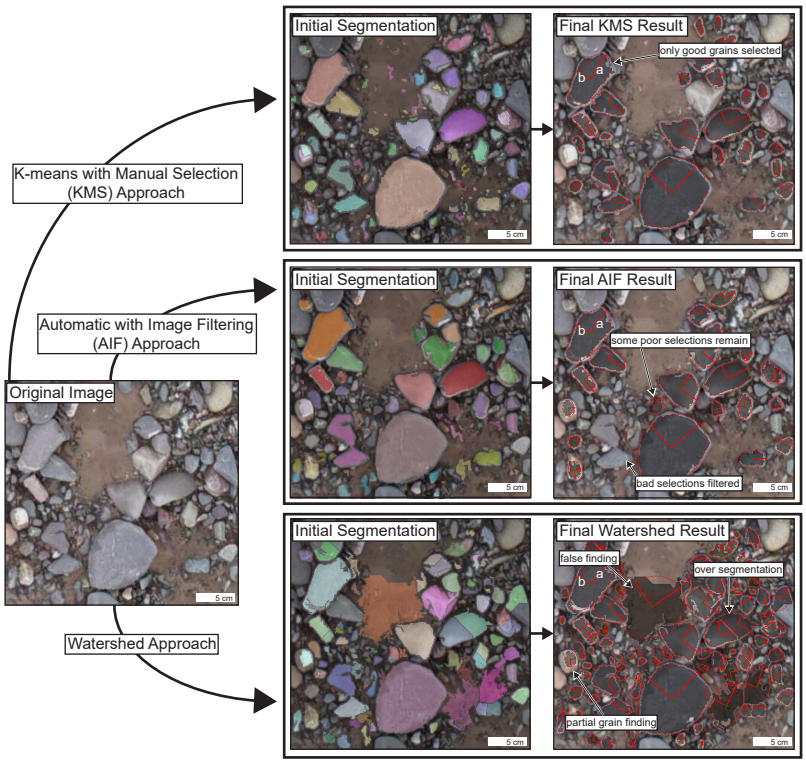


**Figure 4.1** – Difference between our K-means with Manual Selection (KMS) and Automatic with Image Filtering (AIF) approaches versus a fully automated watershed segmentation approach on a gravel image from a high-mountain river. The a- and b-axes of each grain mask are found via an ellipse fit to the same area. Fewer grains are found in the KMS and AIF results, and there is still some misidentification in the case of AIF, but less than in the watershed result.

delineated, thus improving the resulting distribution (Fig. 4.2), with the intention of up-scaling to include many thousand grain measurements over large areas. Despite the selection of fewer grains, Figure 4.2 demonstrates that these represent the true grain size through the close match in distribution with hand-clicked results.

Furthermore, faced with diverse camera models and the rise of SfM-MVS for the generation of georeferenced orthophotos, we wish to explore reasonable and appropriate combinations for covering acre to hectare areas while maintaining accurate grain-size measurement. Fundamentally, our aim for the KMS approach is not in the delineation of a single high-resolution image from a ~1 $m^2$ patch as in previous segmentation work, but rather a method that can cover areas of 10–100 $m^2$ containing complex grain arrangements, despite missing many grains at the patch scale. These semi-automated photo-sieving results can then be used to validate the AIF method at much greater spatial scales ($10^2$–$10^4$ $m^2$), where physical counting is infeasible and previous methods are unreliable or time consuming.

# Algorithm Description

Our methods are similar to previous work by Graham et al. (2005a) and Detert and Weitbrecht (2012), with some key differences. A flow chart of both methods is shown in Figure 4.3 and the processing is presented briefly. We

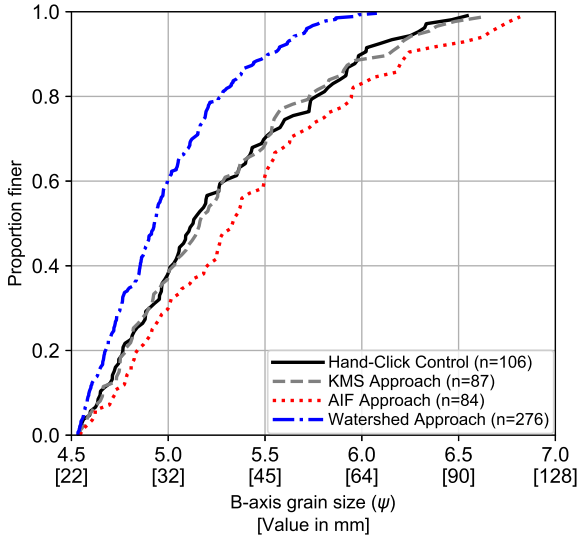


**Figure 4.2** – Watershed segmentation (blue, dashed and dotted line) versus KMS (gray, dashed line) and AIF (red, dotted line) approaches compared with a hand-clicked b-axis grain-size distribution (black line) for a ~1 $m^2$ river patch (S09 in Figure 4.6b). Watershed approach leads to over-segmentation of grains, giving an unreasonable number of clasts (276 versus 106 in the control) and an overly fine grain-size distribution.

direct the interested user to the manual (https://github.com/UP-RS-ESP/PebbleCounts) for a full description of the steps. Our algorithms use 2D image processing in the spatial and spectral domains, which ignores the potential to exploit third height dimensions from irregularly spaced point clouds generated via lidar or SfM-MVS. The reader is directed to Appendix B for our efforts in this regard.

## *PebbleCounts*: K-means with Manual Selection (KMS)

We employ the additional color spaces HSV (hue, saturation, value) and CIELab (Russ, 2002), aside from traditional RGB (red, green, blue) and gray-scale, to enhance differences in the spectral domain separate from lighting. First, the RGB image undergoes strong non-local means denoising (Buades et al., 2011) to smooth intra-granular color difference, interactive gray-scale shadow masking (Otsu, 1979) to separate obvious interstices, and HSV color selection for sand-patch masking (whereby sand is filtered by a narrow, user-selected color mask). The image and shadow/sand edge mask are then windowed for further processing.

At each window, the RGB image undergoes another weaker non-local means denoising, is then converted to CIELab, and the chromaticity bands from this color space undergo bilateral filtering (Tomasi and Manduchi, 1998) to preserve inter-granular edges while further smoothing color. Following this, edge detection on the smoothed, gray-scaled image occurs via a combination of top-hat, Sobel, and Canny methods with feature-AND selections (Russ, 2002), in which an edge is added to the full mask only if it overlaps with a found edge in the previous edge mask, thus piece-wise building an edge map while avoiding lone (i.e., intra-granular) edges (Detert and Weitbrecht, 2012).

After edge detection, our algorithm uses k-means clustering (Lloyd, 1982; Sculley, 2010) to further segment the pebbles. First, the matrix of non-masked pixels is converted into a vector that includes the spectral information at

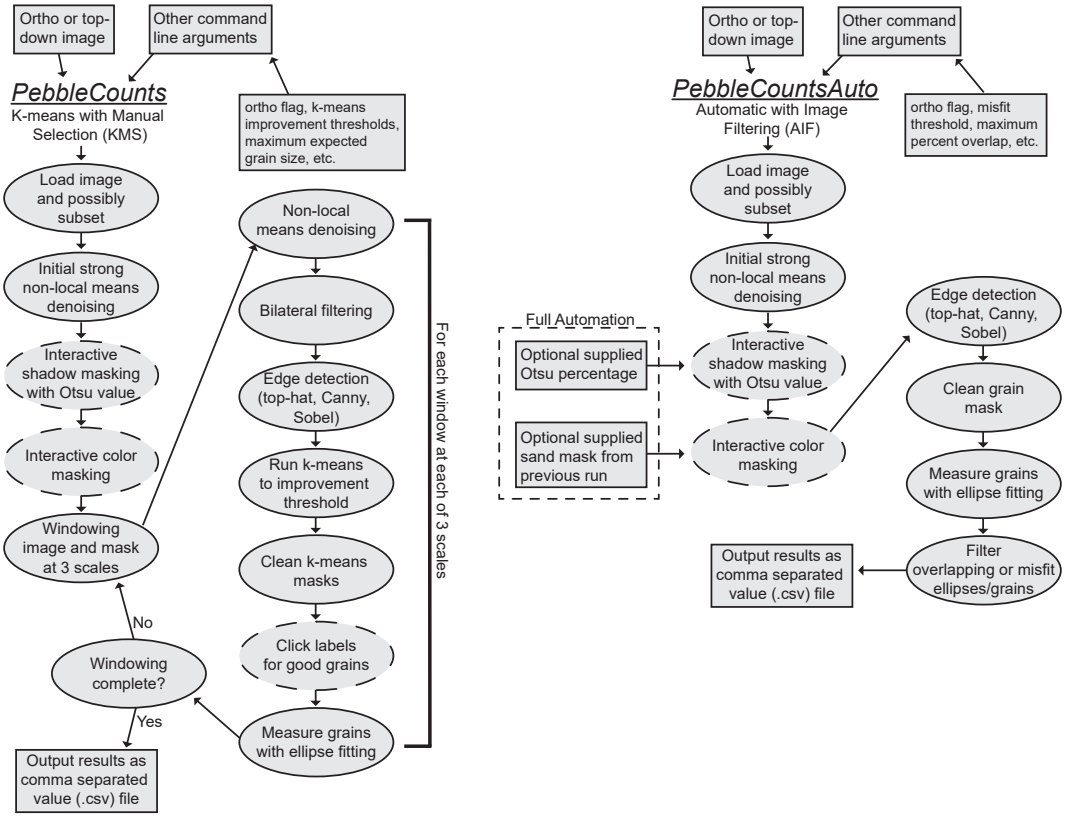


**Figure 4.3** – Flowchart of *PebbleCounts* (left) and *PebbleCountsAuto* (right). The boxes are user supplied input or output from the algorithm. Dashed lines indicate a user input step during processing, either entering and checking values or clicking.

each location. This $N \times 4$ dimensional vector ($N$ being the number of non-masked pixels) includes two spectral observables: the green-red and blue-yellow smoothed chromaticity bands from CIELab; and the two spatial observables: the $x$ and $y$ coordinates of the pixel in image space. To avoid over-segmentation by anisotropic or image-spanning grains, the $x, y$ coordinates are rescaled to 50% of the color, which is also rescaled from 0 to 1.

We attempted using agglomerative Ward hierarchical clustering (Ward, 1963) to further improve results on anisotropic and/or large grains, however, this approach is prohibitively slow on large images, and test results did not show significant improvement. K-means clustering requires a user-supplied number of clusters. Here, we add clusters beginning at 1 and recalculate up to an inertia improvement threshold of 1–10% (user supplied). Resulting labeled masks are cleaned via binary operations and the user is prompted to select the labeled regions that contain full, single grains within a simple pop-up window (Fig. B3).

After selection, the orientation and a- and b-axes of an ellipse fit to the labeled region, shown to accurately approximate grain size (Graham et al., 2005a), are recorded and the grain is added to the final list and the masked region. This processing takes place over three separate scales representing a "burrowing" of the algorithm through the image (from largest to smallest window/grain size). Scales are set by the user supplied longest expected a-axis and image resolution. In contrast to the 46 variables employed by *Basegrain*, *PebbleCounts* has 20 command-line

variable flags — of which 15 exert influence on the results — with most requiring little to no modification (Table B1). Examples of the command-line interface and clicking steps are in the manual and Appendix B.

### *PebbleCountsAuto*: Automatic with Image Filtering (AIF)

This method applies the same initial non-local means denoising and interactive shadow/sand masking, with the option to input user supplied values for full automation. From here, we diverge from the windowing and k-means approach and move directly to edge detection on the entire image using the same top-hat, Canny, and Sobel combination with feature-AND selections.

The resulting mask is then cleaned via binary morphological operations and each label is measured via ellipse fitting. To reduce the misidentified grains, the ellipses are filtered in a three-step chain: (A) Does the centroid fall within another ellipse?; (B) Does the ellipse overlap with any neighboring ellipses above some threshold?; and (C) Is the percent misfit (ellipse area vs. grain-mask area) above some threshold? At each step, an answer of yes leads to the elimination of the grain. The (A) and (B) steps filter grains that have high overlap or are over-segmented, whereas (C) helps filter areas where multiple grains were combined in one mask or a non-grain was identified (e.g., remaining sand patch). Grains passing the test are taken as the final results, with the assumption that misidentified grains are minimal, particularly when up-scaling to large areas and tens-of-thousands of pebbles on high-quality (low-blur) images. The command-line variables for this method are shown in Table B2, and command-line examples can be found in the manual.

We experimented with resampling (over- and under-sampling) the image prior to grain detection to increase smoothing and to improve the detection of larger grains at the cost of measuring fewer smaller grains. The majority of images achieved the best results using the original resolution, though we did find a slight improvement in results using under-sampling on some unsharp images (see Appendix B). The selection of other parameters like the maximum percent misfit is also covered in Appendix B.

# Calibration and Validation I: Controlled Experiment

## Experimental Setup

To test the KMS and AIF approaches on a simple control we arranged three distributions of well-rounded, river pebbles with a-axis sizes from 3–130 mm in semi-overlapping patterns in a $0.5 \times 0.5$ m area (Fig. 4.4). As opposed to most studies that use b-axis lengths to measure the grain-size distribution (Bunte and Abt, 2001), in the experimental setup we use a-axes since it was easier to hand-measure the longest axis of the > 200 grains used. Six size-class bins (3–5, 10–20, 25–35, 40–50, 60–70, and 80–130 mm; all a-axis) were sampled to approximate two log-normal and one bimodal grain-size distribution. The river pebbles used had uniform intra-granular color with minimal striations (i.e., veins), low angularity, and a diverse array of inter-granular colors. Lighting was controlled by overhead fluorescent bulbs and the photos were taken without flash to limit cast shadows.

## Orthomosaic Generation

We tested a Fujifilm X100F model camera with a fixed 23 mm focal length lens and a Sony $\alpha$6000 model with a removable 35 mm fixed length lens. Both had the same advanced photo system type-C (APS-C) sensors (23.6

mm×15.6 mm) and both output photos at 24 megapixels in a 4000×6000-pixel format. Following initial tests, it became clear that the image quality and grain-size results were practically identical for these two cameras, so the results presented are only those for the Fujifilm, as the photo quality was slightly sharper throughout and less distorted at the image corners. To simulate reduced quality, the 24 megapixel Fujifilm picture dimensions were reduced to 75, 50, and 25%, resulting in 13.5, 6, and 1.5 megapixel images at pixel dimensions of 3000×4500, 2000×3000, and 1000×1500, respectively.

For each test setup, we collected ~10 images from ~20°off-nadir (oblique) and at least 4 overhead near-nadir (tilts < 10°) pictures, for 12–16 photos in total. The collection of oblique images aided in removing doming effects from the resulting point clouds (e.g., James and Robson, 2014) and for capturing the pebble edges and sides (Fig. B1). As consumer-grade cameras have square pixels with negligible difference in horizontal and vertical resolution, the image scale can be calculated directly from the camera parameters and camera height with the resolution ($R$) in mm/pixel given by:

$$R = \frac{(S \cdot h)}{(f \cdot I)} \tag{4.1}$$

where $S$ is the sensor height or width in mm, $f$ is the lens focal length in mm, $h$ is the camera height in mm, and $I$ is the image height or width in pixels. $S$ and $I$ should either both be the width, or both be the height of the sensor and image, respectively. This assumes no major distortions within the field of view, which is not valid for oblique imagery, but is negligible for near-nadir photography at close range using non-fisheye lenses. With h=1.55 m, the resulting image resolutions tested from the Fujifilm were 0.26, 0.35, 0.53, and 1.05 mm/pixel by eq. 4.1.

We used the 12–16 photos to generate SfM-MVS orthoimages in *Agisoft Photoscan v.1.4.2* (Agisoft, 2019) — renamed *Agisoft Metashape* in recent versions. This allows rapid output of additional information including point clouds, digital elevation models (DEMs), and the undistorted orthomosaics, with resolution recorded in the image metadata for direct input into *PebbleCounts* and *PebbleCountsAuto*. Detailed *Agisoft* processing steps are provided in Appendix B.

## Comparison Metrics

For the simple, controlled experiment, with relatively coarse grain-size bins, it is not appropriate to compare percentiles (e.g., $D_{50}$) or to run Kolmogrov-Smirnov (KS) tests and measure the difference in distributions between the AIF or KMS and control grain-size distributions. Instead, we compared the counts in each bin between the control and algorithm and visually assessed the matching of the grain-size distributions. This provides a reasonable baseline for checking the performance of the algorithm in a highly controlled setting.

## Results I: Controlled Experiment

For each of the three 150–200 clast arrangements, the KMS *PebbleCounts* run time was ~7 minutes on a laptop with 16 GB RAM and 2 cores (Intel i7-6650U 2.20 GHz) and no GPU, whereas the AIF *PebbleCountsAuto* run time was ~1 minute. Both a single near-nadir image and the combined orthomosaic were used, but the results were entirely consistent aside from some inter-run variability in the KMS approach caused by the non-unique solution

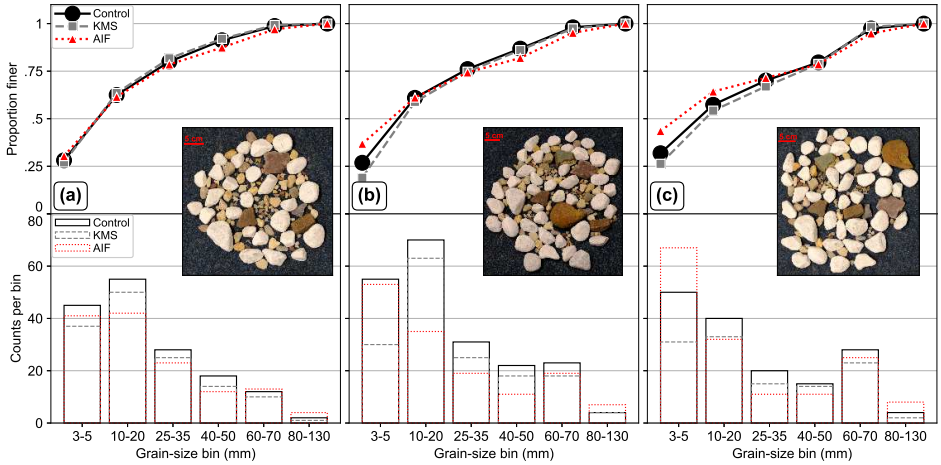


**Figure 4.4** – Result of KMS (gray, dashed lines) and AIF (red, dotted lines) on the three experimental lab setups (a-c) with known grain inputs in six size classes (black line), measured as the grain a-axis. (a) Log-normal, (b) log-normal with increased number of all classes, including fines, and (c) skewed bimodal with increased number of coarser grains. Bottom row shows the counts per bin and the top row shows the resulting grain-size distribution. The images are 0.26 mm/pixel (24 megapixel).

of k-means clustering. Given this consistency, we only present the results from the single near-nadir images. Furthermore, the use of only 4 overlapping near-nadir photos also generated the same results, albeit in about $1/6^{th}$ the *Agisoft* orthomosaic processing time of using all 12–16 photos (~10 minutes versus ~1 hour on the same laptop).

Across all three distributions, the KMS approach consistently undercounts the number of clasts in each a-axis bin (Fig. 4.4). However, and in agreement with previous research (Graham et al., 2010), this undercounting is uniformly distributed and thus the grain-size distributions do not show notable differences between the algorithm and control. For the two arrangements with increased fine (3–5 mm) and coarse (60–130 mm) pebbles (Fig. 4.4b,c), the undercounting is stronger at the finer end of the distribution leading to a slight underestimation of the grain-size distribution by the KMS approach in this region. This is caused partially by the user missing more of the smaller grains (of which there are exponentially more), some smaller grains being partially hidden by the larger, and also by the smallest grains being only a few pixels in area and thus eliminated during mask-cleaning steps, or not captured at all. On the other hand, the AIF approach tends to overcount the fine pebbles, leading to overestimation of the grain-size distribution, because many small non-grain areas remaining in the masked image are automatically selected in the final result, rather than ignored as in the KMS approach.

As we reduced the resolution from 0.26–1.05 mm/pixel, the reduction in the finest size class increased dramatically for the KMS approach (Fig. 4.5). At the lowest resolution tested (1.5 megapixel), this undercounting leads to severe discrepancies in the grain-size distribution curve. As the resolution degrades it becomes more difficult to discern rocks in the smallest size class (3–5 mm), which correspond to an a-axis grain size of 12–19, 9–14, 6–9, and 3–5 pixels for the 24, 13.5, 6, and 1.5 megapixel resolution, respectively, indicating the necessity of a limiting lower measurement factor (e.g., Graham et al., 2005a).

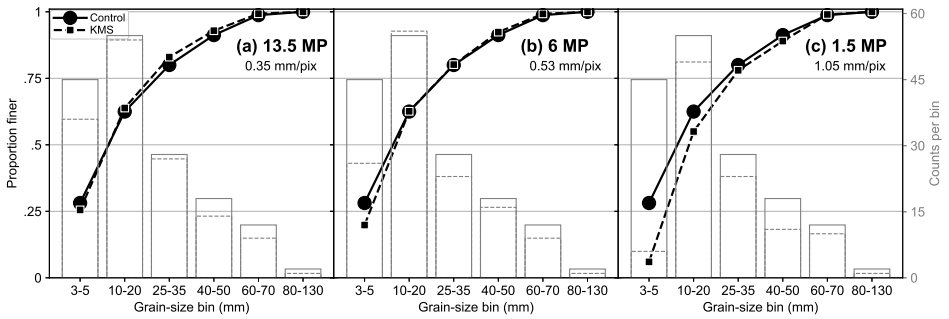


**Figure 4.5** – Results of reducing the image dimensions to (a) 75% (13.5 megapixel), (b) 50% (6 megapixel), and (c) 25% (1.5 megapixel) and re-running the KMS approach on the distribution in Figure 4.4a. Control is shown as black (left y-axis) and gray (right y-axis) solid lines and KMS as the dashed lines.

# Calibration and Validation II: Field Surveys

## Field Setting

Having established the algorithms on control data, we sought to evaluate the performance on complex, natural photos. Field data provides the real-world application and detailed uncertainty analysis most useful for researchers seeking to apply the methods to their own sites. For this we turned to photo surveys carried out on gravel-bed river cross sections of the foreland and topographic transition zone of the northwestern Argentine Andes (Fig. 4.6). This is an area of strong precipitation, topographic, and environmental gradients, and the dynamic rivers surveyed are capable of transporting enormous quantities of sand, gravel, and boulders of various lithology (Bookhagen and Strecker, 2012; Purinton and Bookhagen, 2018). Catchment-average erosion rates from the area, based on cosmogenic nuclide inventories, suggest rates on the order of 0.6–1 mm/yr (Bookhagen and Strecker, 2012), with large variability during the Pleistocene and Holocene (Tofelde et al., 2017). The region is frequently affected by extreme hydrometeorologic events that lead to flooding and drainage-pattern rearrangement (Castino et al., 2016a, 2017).

## Orthomosaic Generation

All cross-section surveys were collected using the Sony $\alpha$6000 camera model at 24 megapixel resolution, and survey sizes ranged from ~1000–5000 $m^2$. In this case, the standard zoom lens delivered with the camera was used at the shortest focal length of 16 mm to maximize the field of view. Also, to help cover the large survey sites, the camera was affixed to the end of a pole with a remote control trigger, allowing overhead shots to be collected from a height of 4.5–5 m (Fig. 4.7), giving a ground resolution of approximately 1.1–1.2 mm/pixel by eq. 4.1. UAV flights have proven difficult in the windy conditions experienced in these valleys, but flights at 20–30 m heights with the 12 megapixel camera provided on the DJI Mavic and Phantom models (focal lengths of 3.6–4.3 mm, sensor dimensions of 6.17×4.55 mm, and image dimensions of 4000×3000 pixels) would result in image resolutions of ~7–13 mm/pixel, and are thus inadequate for delineating cm-scale pebbles.

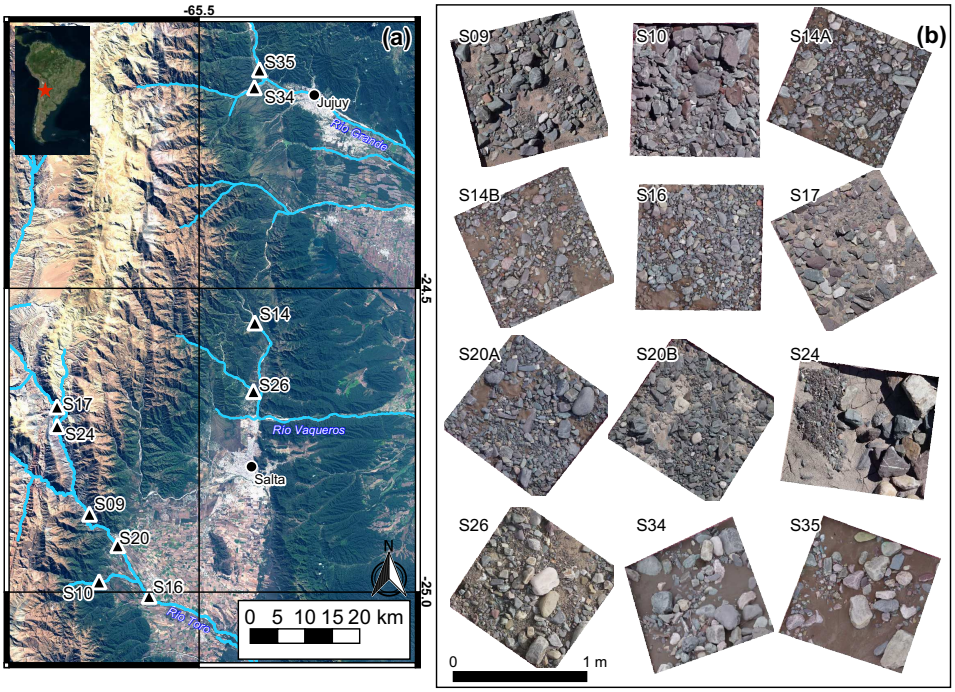


**Figure 4.6** – (a) Field cross-section survey sites (black triangles) in NW Argentina from three gravel-bed rivers (Toro, Vaqueros, and Grande) and their tributaries, draining from the sparsely vegetated mountains in the west towards the verdant foreland and city centers of Salta and Jujuy in the east. The Landsat 8 RGB composite satellite image (using bands 2, 3, and 4) from 12 June 2017 shows the climatic transition from wet foreland to dry mountains, demarcated by the green-brown transition zone, running approximately north-south, corresponding to vegetation changes. (b) Detailed view of the 12 × ~1 $m^2$ orthomosaic clips from each of the field sites with average resolution of 1.16 mm/pixel.

To generate georeferenced orthomosaics that could be tiled and passed directly to *PebbleCounts* and *PebbleCountsAuto*, survey sites on the dry river-bed were laid out with on average 18 coded targets (with a range of 10–24) and the position of each was measured with a differential GPS (Fig. 4.7). Kinematic post-processing with a permanent base station < 100 km away at the Universidad Nacional de Salta (UNSA) in Salta, Argentina, led to cm accuracy of XYZ target locations. The site was traversed in a cross-hatched pattern with a photo captured every 2–3 paces, so that each location appeared in ~9 near-nadir pictures from slightly different angles. We refer to the images as near-nadir, rather than nadir, due to the fact that during mast photo collection some unintentional tilting of the camera (< 10°) occurred. These near-nadir photos aided in removing doming effects, but did not allow us to capture the sides of pebbles as in the oblique images taken in the experimental setup (Fig. B1). Capturing oblique images of every patch in the field sites would require infeasible amounts of time and processing power.

*Agisoft* processing was similar to that described for the experiment, with some key differences (see Appendix B). Given the volume of photos (600–1300 per site), the sites were processed automatically using the Python API for *Agisoft*, with processing times consistently over 10 hours on an 80 core, 500 GB RAM server making use of 1 GPU NVIDIA Tesla K80 unit for some of the steps (e.g., dense matching).

**Figure 4.7** – Sony $\alpha$6000 24 megapixel camera affixed to mast for photo collection at a height of 4.5–5 m (left) and differential GPS measurement of coded targets (right).

From 10 of our full survey sites over three different river systems we selected 12 $\times$ ~1 $m^2$ patches to clip out of the full orthomosaics and evaluate using the KMS and AIF approaches. The final resolution of these 12 GeoTiff orthoimages matched the theoretical value from eq. 4.1, with an average of 1.16 mm/pixel and range of 1.08–1.24 mm/pixel (standard deviation of 0.05 mm/pixel). The patches (Fig. 4.6b) include variable amounts of sand and a large range of grain sizes, packing arrangements, and shadowing. From one site (S14A) there were handheld images available for the same selected patch from the same Sony $\alpha$6000 camera zoomed to 20 mm focal length and taken from a height of ~1.5 m, allowing for the generation of a complementary orthomosaic at 0.32 mm/pixel resolution.

## Comparison Metrics

For control data from the field we return to b-axis measurements (rather than a-axes as in the lab). In each patch, the b-axes of all grains visible to the naked eye were manually digitized. This generated a 5490 pebble control dataset across all 12 mast-surveyed sites. For the lone handheld patch at 0.32 mm/pixel, the control data was 1726 pebbles versus 621 from the same patch at the 1.12 mm/pixel mast resolution, as smaller grains could be manually measured on the image at a 4-times improved resolution.

The use of continuous control data, as opposed to discrete bins in the lab experiment, allows a more detailed investigation of the performance of both approaches, including biases and their correction. B-axis measurements of overlapping control and KMS grains were compared to look for sizing bias. This was followed by a search for the lower truncation limit (the lower cutoff in b-axis length in pixels that grains are reliably measured at) of the algorithm, also using the KMS results. For parts of the analysis, the size data were converted to the typical $\psi$ scale ($\psi = -\phi = log_2(mm)$) of grain-size measurement of coarse river sediments. This allows direct comparison of statistical results with other studies (e.g., Graham et al., 2005b)

We compared the grain-size distributions from the KMS and AIF approaches with the control using a two sample KS-test to check the null hypothesis that the two samples are drawn from the same distribution. Because sample sizes were at times small, leading to erroneous KS-test results, we also devised a second metric of grain-size distribution comparison. Similar to the KS-test, which uses the maximum distance between the cumulative distribution functions (grain-size distributions), our metric interpolates both distributions to the same lengths in 0.1 $\psi$ steps and then sums the difference between the re-interpolated curve to give an approximate integral of the difference between the two grain-size distributions (AIF or KMS minus the control), which we term $A_{diff}$. Here, an $A_{diff}$ value close to 0 indicates good matching, and positive or negative values indicate underestimation or overestimation, respectively.

We also examined the performance of some key percentiles ($D_{5,16,25,50,75,84,95}$). The metrics for comparison of control ($P_C$) and KMS or AIF ($P_P$) percentiles are consistent with other studies (Sime and Ferguson, 2003; Graham et al., 2005b, 2010). These are the mean ($m = \frac{1}{n} \cdot \Sigma(P_P - P_C)$), the mean squared ($ms = \frac{1}{n} \cdot \Sigma(P_P - P_C)^2$), and the irreducible random error ($e = \sqrt{ms - m^2}$). The bias of *PebbleCounts* is quantified by $m$, and $e$ measures the scatter or precision after bias correction (Sime and Ferguson, 2003).

# Results II: Field Surveys

The KMS *PebbleCounts* approach took ~10 minutes per 1 $m^2$ orthomosaic clip at 1.16 mm/pixel resolution, depending on the number of grains, and particularly the number of finer grains, present. Run time for the AIF *PebbleCountsAuto* approach was typically ~2 minute per site. All run times refer to the same laptop with 16 GB RAM and 2 cores (Intel i7-6650U 2.20 GHz) and no GPU. For the 0.32 mm/pixel image the processing for KMS took ~45 minutes, as there were more fine grains to be identified (given the log-normal distribution) and so the clicking took exponentially longer, and the AIF took ~20 minutes given the longer time spent filtering the large number of grains. We note that the use of a GPU for the filtering steps will significantly improve processing time. Importantly, these run times refer to the use of no lower truncation value, which leads to much longer processing time.

An aggregation and coarse binning of all b-axes in the control versus KMS and AIF data for the coarser imagery are presented in Figure 4.8. There is obvious undercounting in these data from the KMS results, similar to the experimental setup, and here it causes a significant discrepancy in the grain-size distribution curves. Whereas the manual clicking found over 1000 grains in the smallest classes (1–2 and 2–3 $\psi$), the KMS approach found none in the smallest and only ~100 in the second smallest. This skews the percentiles to the higher grain sizes, and thus overestimates them significantly. In opposition to this, but again in agreement with the experimental setup, the AIF results display significant overcounting at the finer sizes as many non-grains are identified, particularly when the algorithm is run with no lower truncation.

The skewed results from both the KMS and AIF approaches warrant detailed analysis of the algorithms' deficiencies and grain-size distribution corrections. To begin, we examined the performance of *PebbleCounts* on grains manually digitized and the same grains selected during clicking in the KMS approach on the coarser imagery (Fig. 4.9). There is only a slight negative bias across all grain sizes, indicating underestimation of individual grains by *PebbleCounts*, however, this median shift varies with no apparent pattern and is likely caused by uncertainties in the manual b-axis digitization of thousands of grains. For instance, digitization with b-axis vector lines can achieve sub-

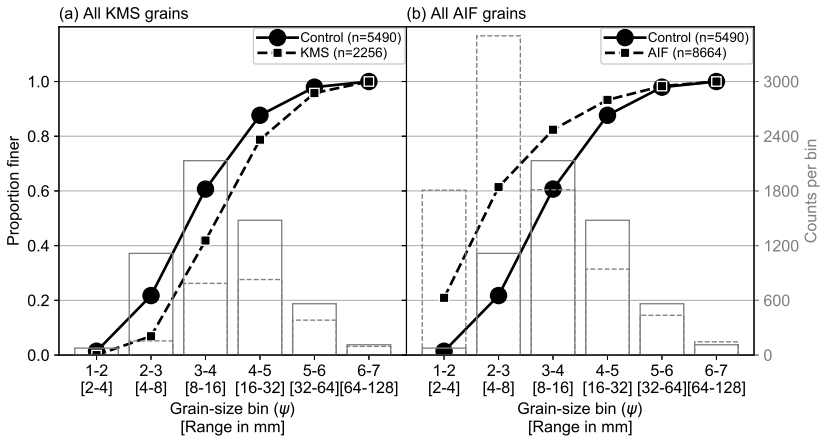


**Figure 4.8** – Comparison of (a) KMS and (b) AIF at the 12 field sites all aggregated and coarsely binned. Control is shown as black (left y-axis) and gray (right y-axis) solid lines and KMS and AIF as the dashed lines.

pixel accuracy compared to the raster processing of *PebbleCounts*. The AIF approach measures grains identically to the KMS method and thus has the same misfit errors on correctly identified grains. From this we conclude that the algorithm is effective on a grain-by-grain basis and the skewing of the grain-size distributions are instead caused by sampling errors related to the image resolution and ability to find small grains (see Figure 4.5).

The undercounting error can be explored on the full distribution of pebbles by gradually increasing the lower truncation value and assessing the error in percentiles versus the control data at each step (Fig. 4.10). As truncation is increased, the median percentile error decreases rapidly up to an inflecting value — manually chosen from the graph as a significant local minimum where the median difference is near 0 mm. Truncating the KMS distributions at a minimum b-axis length of 23 mm (rounded to 20 pixels) improves the results significantly for the 1.16 mm/pixel imagery taken from the mast. Beyond this truncation, there is limited improvement. Regarding the 0.32 mm/pixel image, the 20-pixel (6.5 mm) truncation also results in a median difference near 0 mm, with subsequent truncation values leading to only ~0.5 mm improvements. Supplying these truncation values directly to the KMS *PebbleCounts* tool results in reduced processing time to ~5 minutes for the coarser imagery and ~15 minutes for the finer, as many small grains were then ignored and left out of the clicking mask.

The same analysis for the AIF approach is complicated by the large number of false grains found and the extreme overcounting of fine grains. Given this, we instead make the assumption that the similarity of the two methods, particularly in the edge detection and ellipse fitting steps, leads to similar errors in both. Therefore, we assume the same 20-pixel truncation. For the AIF *PebbleCountsAuto* tool, processing times with the 20-pixel truncation reduced to < 1 minute and ~3 minutes for the coarse and fine images, respectively.

The combined results before and after lower truncation for the coarser (~1.16 mm/pixel) imagery taken from the mast surveys is shown in Figure 4.11. Without any lower truncation, the AIF tool results in significant overcounting and grain-size distribution underestimation with a high $A_{diff} > 8$. The KMS tool instead shows undercounting and grain-size distribution overestimation with a low $A_{diff} < -4$. Both have KS-test $p$-values < 0.0001. When we apply a 20-pixel truncation, both the AIF and KMS approaches achieve $A_{diff}$ values near or below $-1$, with the manual

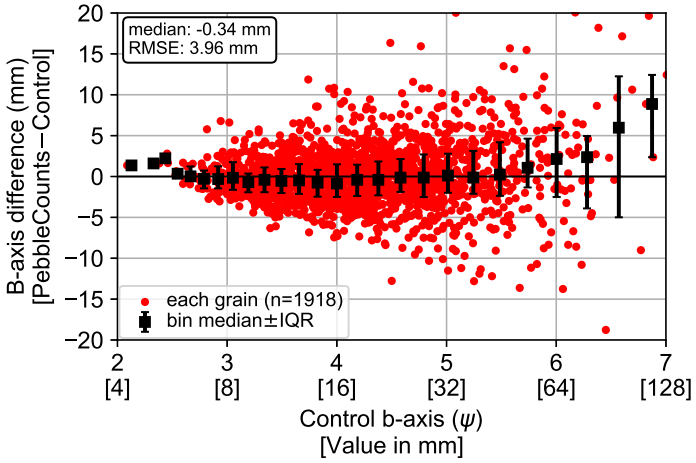


**Figure 4.9** – Measurement error of *PebbleCounts* (here the KMS results) versus control on a grain-by-grain basis for overlapping grains in the coarser (1.16 mm/pixel) imagery. There is an overall median shift, but the binned medians do not display a consistent pattern.

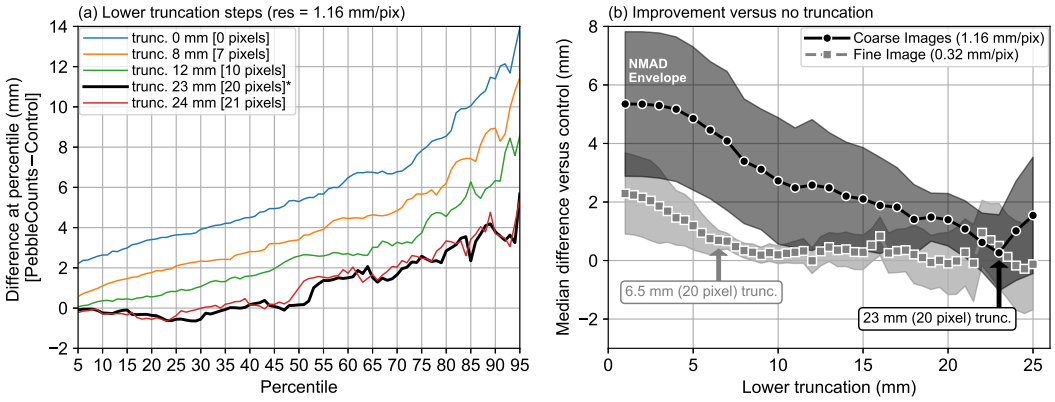


**Figure 4.10** – (a) Error in each percentile (5–95) as lower truncation value is increased in 1 mm steps for the 1.16 mm/pixel imagery. Only a few steps are plotted for clarity. (b) The median difference in percentiles compared with the control versus the lower truncation value, with the normalized median absolute difference (NMAD) shown as the error envelope (Höhle and Höhle, 2009). From this analysis, we select a lower truncation of 20 pixels. The analysis in (a) was repeated for the finer image (with 0.5 mm truncation steps) to get the gray squares line in (b), and is not shown here.

KMS approach performing best and achieving a high KS-test $p$-value of 0.2398. The AIF approach retains a low $p$ of 0.0008 with a ~0.1–0.2 $\psi$ bias towards coarser values in the upper portion of the grain-size distribution (> $D_{50}$).

In Appendix B (Fig. B7), we show the 20-pixel truncated KMS and AIF results on a site-by-site basis. For the KMS approach, following truncation 11 sites have $p$-values > 0.1 and one site (S16) has $p$=0.0971. $A_{diff}$ values are also near 0 indicating close matching of the grain-size distributions, aside from S24 and S34, which both show large discrepancies. The AIF results in Figure B7 follow a similar trend to the KMS results, but there is a bias towards coarser values, with many $A_{diff}$ values < −1, and generally poorer results compared with the KMS approach, with

grain-size distributions being overestimated by ~0.1–0.2 $\psi$. In the KMS results, despite a high $p$-value, S24 demonstrates a stronger bias in the grain-size distribution towards coarser grains (up to 0.5 $\psi$ discrepancy), as indicated by the high $A_{diff}$ value of $-1.36$. Here, the KS-test pass is likely caused by the small sample size remaining after truncation ($n$=24), the least of any site. The poor performance of S24 was expected given the large size range with many sub-cm pebbles and a few large boulders, strong cast shadows from the large grains, and intra-granular edges on angular boulders with quartz veins (see Figure 4.6b). Importantly, S24 is the only site not from a major river stem, but rather from a debris-flow fan draining a small tributary catchment in the Quebrada del Toro. S34 also had a high $A_{diff}$=$-2.11$. In this case, poor performance is due to significant blurriness of this image, and again a small sample size ($n$=47).

We also compared the individual percentiles of interest to assess the bias and accuracy of truncated results (Fig. 4.12). For the KMS approach, the bias ($m$) is 0.06 $\psi$ with a precision ($e$) of 0.13 $\psi$. Excluding S24 and S34, $m$ and $e$ drop to 0.03 and 0.09 $\psi$, respectively. The AIF results have higher $m$ and $e$ values of 0.15 and 0.17 $\psi$, respectively, which are reduced to 0.13 and 0.15 $\psi$ following exclusion of the same S24 and S34 sites, in addition to the S10 site, which was also somewhat blurry and with relatively few grains. For the AIF percentiles, we chose to include S16 despite large overestimation at higher percentiles (Fig. B7), as this was a sharp image with a relatively large sample size. The high uncertainties from this scene likely require some adjustment of the edge-detection variables (see Appendix B) for improved segmentation, but the results presented are realistic for fast processing using the AIF method, with the caveat of higher expected uncertainties.

The uncertainties in Figure 4.12 are average values, and the inset plots also demonstrate the increasing uncertainty of larger percentiles. The maximum uncertainty for both at $D_{95}$ is $m$=0.08 $\psi$ and $e$=0.07 $\psi$ for the KMS result and $m$=0.35 $\psi$ and $e$=0.2 $\psi$ for the AIF result. Importantly, since the $\psi$ scale is logarithmic, the larger errors at higher percentiles correspond to similar percentage misfits as lower errors at smaller percentiles (e.g., 0.2 $\psi$ precision at a grain size of 6.5 $\psi$ (91 mm) is a 13–15% misfit, whereas, a 0.01 $\psi$ precision at 4.5 $\psi$ (23 mm) is a 4–10% misfit).

As a final test for the KMS and AIF approaches, we turn towards our handheld imagery taken from S14A with a 4-times improved resolution of 0.32 mm/pixel (Fig. 4.13). We only show the 20-pixel truncated results, which displayed high KS-test $p$-values > 0.2 and $A_{diff}$ close to 0 in both cases, with the AIF approach slightly underestimating ($A_{diff}$=0.6) and KMS slightly overestimating ($A_{diff}$=$-0.77$). For the KMS approach $m$ and $e$ are 0.07 and 0.05 $\psi$, respectively, and $-0.06$ and 0.05 $\psi$ for AIF.

## Caveat of *PebbleCountsAuto* AIF

The promising results of the AIF approach shown in Figures 4.11–4.13 come with some consideration of the grain-by-grain accuracy. In Figure 4.14, we analyze the percentage of grains found in the AIF approach that have a corresponding grain in either the hand-clicked control (based on a 6-mm buffer of the b-axis line) or the KMS results (based on a 6-mm centroid buffer). From this subset of grains, we consider the AIF grain to be a matching (or correct) result if the b-axis difference between it and the nearby "good" grain (from the control or KMS) is < 1 cm. From this we see that in the best-case scenario the percentage of correct grains identified by the AIF approach is only 70%, from the handheld 0.32 mm/pixel image. A number of sites (S10, S16, S20B, S24, S34, and S35) have <

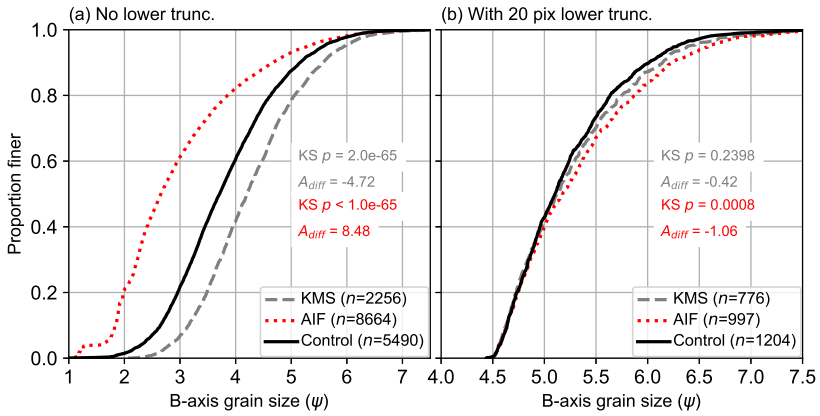


**Figure 4.11** – Results from hand-clicked control (black line), KMS *PebbleCounts* (gray, dashed line), and AIF *PebbleCountsAuto* (red, dotted line) with the initial non-truncated run (a) and the 20-pixel truncated run (b). In corresponding colors are the $p$-value results of a KS-test and the $A_{diff}$ approximate integral between the curves for each approach versus the control data. The legend indicates the number of grains ($n$) making up each curve. Note the reduction in x-axis scale between the columns, where the right, truncated distributions are plotted on a narrower range to emphasize the remaining discrepancies. The curves separated by site (Fig. 4.6b) are shown in Figure B7

50% matched grains. The two poorly performing sites (S24 with grain complexity and S34 with image blur) both demonstrate the lowest accuracy with < 40% matches. Notably, despite a significant number of false positives in the results, when comparing the overall grain-size distributions (Fig. 4.11), and on a site-by-site basis (Fig. B7), the distribution of the AIF results matches the hand-clicked control well. The errors associated with the AIF method are demonstrated in Appendix B.

# Discussion

In this study we developed two new methods for grain-size measurement with low uncertainties and the potential to deliver full grain-size distributions from complex images of high-energy mountain rivers. Our open-source Python-based algorithms perform equally well to other image segmentation tools, but can be applied more quickly over larger areas surveyed by the SfM-MVS workflow we present. Critical to success is the application of a strict lower cutoff, which limits the minimum measurable b-axis grain size to 20-times the pixel resolution. The automated version of the algorithm delivers less accurate measurements, but these can be limited by using low-blur, higher resolution imagery. We focus our discussion on the comparison of our approach with similar work, the effect of the lower truncation on grain-size distribution estimates, and practical guidelines for acquiring imagery and applying *PebbleCounts*, including the application of UAV surveys.

## Performance

For comparison of our algorithms to previous work, we do not consider errors reported in studies using texture-based measurements (e.g., Woodget et al., 2018), since these are based on correlative relationships rather than physical measurement of each grain. Texture methods work well for homogeneous pebble arrangements in lower-energy

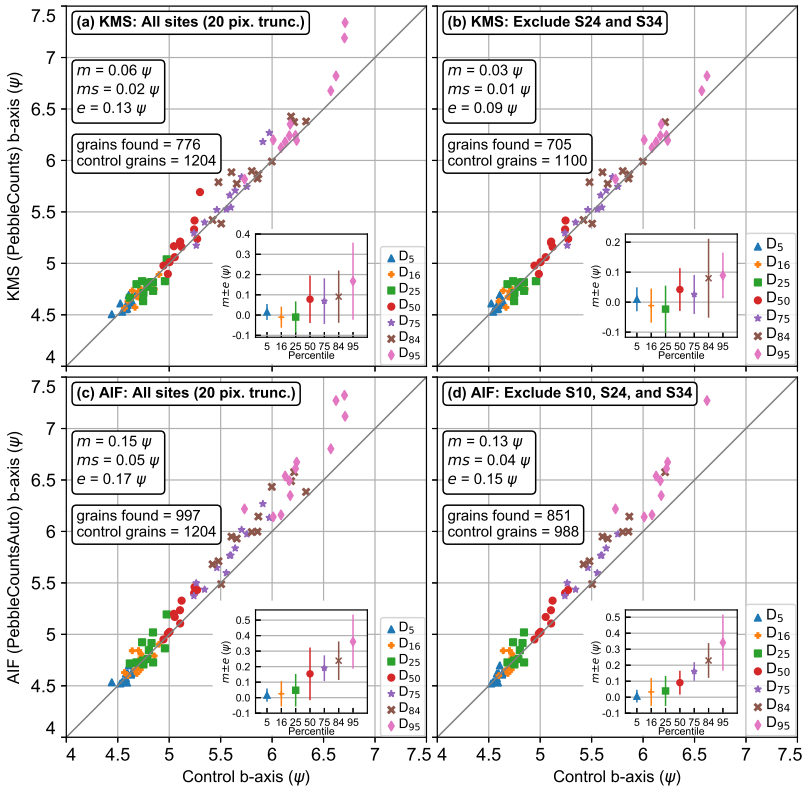


**Figure 4.12** – Comparing the key b-axis percentiles across all 12 field sites and between the KMS and AIF approaches with the 20-pixel truncation applied. (a) All 12 sites from KMS, (b) KMS improvement when excluding S24 and S34, (c) all 12 sites from AIF, and (d) AIF improvement when excluding S10, S24, and S34. For the main plot, each data point is a percentile value from a single site and the 1:1 relationship is the gray diagonal. The mean ($m$), mean squared ($ms$), and irreducible ($e$) errors are shown for each plot, taken as the average of all 7 percentile errors across the 9–12 sites plotted. The $m$ and $e$ are separately plotted for each percentile in the inset plot. The number of grains in the control ("control grains") and KMS or AIF results ("grains found") are also indicated. The individual site curves where these data points originate are shown in Figure B7.

settings, but high-energy mountain rivers with heterogeneous pebble arrangements and large ranges in sizes require segmentation approaches. Similar to other image segmentation methods, the KMS *PebbleCounts* approach undercounts grain sizes in each respective size class (Graham et al., 2010). This undercounting does not undermine the resulting grain-size distributions and associated percentile estimates, so long as an appropriate lower truncation is defined. This cutoff was found to be 20 pixels (compare to 23 pixels found by Graham et al. (2005a)) in b-axis length (Fig. 4.10), which explains the degradation in 3–5 mm counting in the reduced resolution lab images (Fig. 4.5), where the smallest pebbles were only a few pixels in size as resolution was decreased.

As shown in Figure 4.12, when we apply this cutoff and exclude poorly performing images we find an average $m$ (bias) and $e$ (precision) of 0.03 and 0.09 $\psi$, respectively, for the ~1.16 mm/pixel imagery and 0.07 and 0.05 $\psi$ for the 0.32 mm/pixel image. For the AIF approach these values are 0.13 and 0.15 $\psi$ for the ~1.16 mm/pixel imagery and $-0.06$ and 0.05 $\psi$ for the 0.32 mm/pixel image. These are averages, which actually increase at higher percentiles in

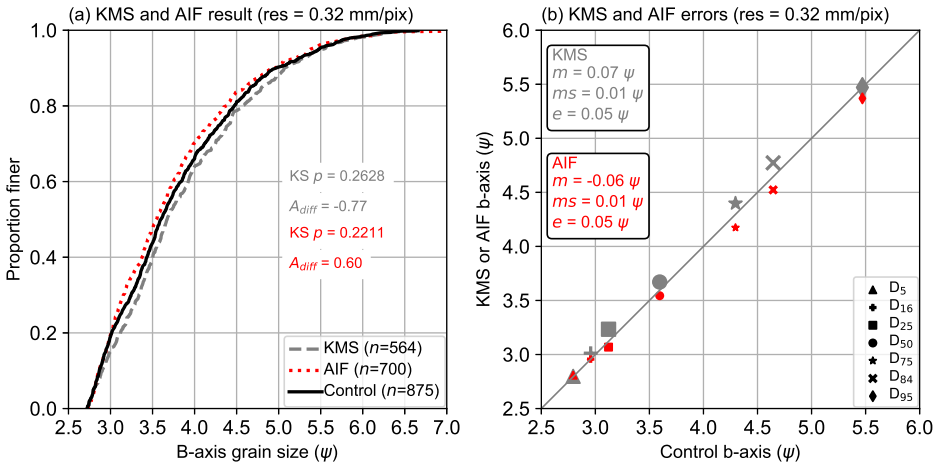


**Figure 4.13** – (a) Results from hand-clicked control (black line), KMS *PebbleCounts* (gray, dashed line), and AIF *PebbleCountsAuto* (red, dotted line) from the 20-pixel truncated run on the 0.32 mm/pixel handheld imagery. In corresponding colors are the *p*-value results of a KS-test and the $A_{diff}$ approximate integral between the curves for each approach versus the control data. (b) Percentile comparison for both methods with KMS in gray and AIF in red, with inset box showing the uncertainties for each in the corresponding color.

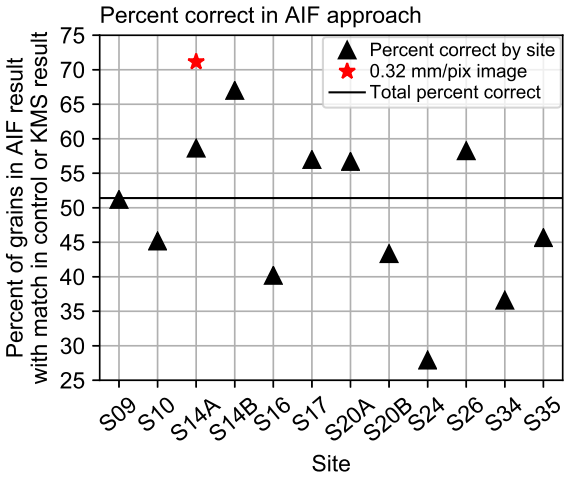


**Figure 4.14** – Percentage of grains from AIF results with a matching grain in either the hand-clicked control or in the KMS result. A match is defined as a grain within 5 pixels of the hand-clicked line or the KMS grain centroid for the 1.16 mm/pixel imagery, or within 20 pixels for the 0.32 mm/pixel image (corresponding in both cases to a distance of ~6 mm), and with a 1 cm maximum b-axis difference between the AIF grain and the match. The total percent correct, taken across all black triangles, is 51%.

agreement with other image segmentation methods (e.g., Sime and Ferguson, 2003). We thus suggest higher error budgets at higher percentiles.

As demonstrated in Figures 4.14 and S8, there are significant inaccuracies associated with the AIF approach. The errors associated with the AIF approach can be limited when applied to high-quality (low-blur) ~1 mm/pixel

**Table 4.1** – Comparison of *PebbleCounts* and *PebbleCountsAuto* results with other segmentation-based pebble counting studies.

| **Study / Technique** | **Bias ($\psi$)** | **Precision ($\psi$)** |
|---|---|---|
| This Study / K-means with Manual Selection (KMS) | 0.03–0.07 | 0.05–0.09 |
| This Study / Automatic with Image Filtering (AIF) | −0.06–0.13 | 0.05–0.15 |
| Butler et al. (2001) / Custom watershed segmentation | 0.13–0.33[a] | - |
| Sime and Ferguson (2003) / Custom watershed segmentation | 0.14[b] | 0.22[b] |
| Graham et al. (2005b) / Custom watershed segmentation | 0.007–0.03 | 0.07–0.09 |
| Westoby et al. (2015)[c] / *Basegrain* (Detert and Weitbrecht, 2012) | 0.16–0.82[d] | 0.33–1.99[d] |

[a] Taken from only three percentiles ($D_{50,84,95}$).
[b] Corrected value presented by Graham et al. (2005b).
[c] Comparison made in mm, converted to $\psi$ units here.
[d] Large spread caused by significant disagreement at higher percentiles.

resolution imagery, with better results possible on < 0.5 mm/pixel imagery. Ultimately, the uncertainties are highly dependent on the input image quality and complexity (range in grain size, angularity, intra-granular variability) and providing blanket estimates is less useful than end-users applying the KMS tool to a subset of images to validate the results of the AIF approach.

In spite of this caveat, our bias and precision values of −0.06–0.15 $\psi$ are on the low end of previously published errors from similar techniques (Table 4.1). To our knowledge, the only study to compare *Basegrain* results to control data by Westoby et al. (2015), makes comparisons in mm rather than $\psi$ units. Since the $\psi$ scale is logarithmic, in our study the error in mm increases with $\psi$ from ∼0.8 mm uncertainty at 4.5 $\psi$ (23 mm) to ∼7 mm uncertainty at 6.5 $\psi$ (91 mm) for the ∼1.16 mm/pixel imagery in the KMS case. Westoby et al. (2015) report even greater bias and lower precision from *Basegrain*, with errors also increasing in magnitude at higher percentiles. We emphasize that the previous image segmentation techniques discussed all rely on watershed segmentation, whereas neither of our algorithms use this step for the reasons demonstrated in Figures 4.1 and 4.2.

# Lower Truncation

The issue of lower truncation on grain-size distributions and percentile estimates has received much attention in the literature (e.g., Fripp and Diplas, 1993; Rice and Church, 1996; Bunte and Abt, 2001; Graham et al., 2010). Previously, field geomorphologists were interested in all grains above 8–16 mm, simply because smaller grains were difficult to manually identify and thus underrepresented in the results (e.g., Fripp and Diplas, 1993; Rice and Church, 1998). Previous work suggests that truncation at the finer end of the distribution primarily increases the lower percentiles, while having less effect on the large (> $D_{50}$) percentiles (Bunte and Abt, 2001). We find significant shifts in all percentiles of > 0.5 $\psi$ when applying a 20-pixel truncation. Graham et al. (2010) report truncation errors of < 0.3 $\psi$ for all percentiles in 1, 3, and 5 $\psi$ truncated distributions. Their better results at lower percentiles are likely because the data were collected manually grid-by-number style in the field with the ability to include smaller grain sizes. The measurement resolution presents the ultimate control on how accurately grain-size percentiles can

be measured. The purpose of the KMS and AIF approaches introduced here is in acquiring grain-size distributions from a subset of the full grain-size range present in the river, namely the subset with > 20-pixel b-axis length in image resolution.

## Image Acquisition

Ideally, collecting 9+ near-nadir images/$m^2$ (as in our field surveys) or collecting an approximately 1:2 (or greater) ratio of near-nadir to oblique imagery (as in our experiments with point-cloud data dimensions; see Appendix B), leads to the highest quality point-cloud results in *Agisoft*. Higher quality point clouds, in turn, lead to less distortion errors during orthorectification and higher quality orthomosaics. Due to the textured nature of gravel images, we attained comparable results in reduced time using only 4 overlapping near-nadir images/$m^2$ in the lab setting. In any case, high overlap of ~80% between images is recommended to ensure the best results. Where a user desires accurate and dense point-cloud data in addition to the 2D orthomosaics, it is recommended that (many) more images closer to the surface be collected (e.g., Verma and Bourke, 2019) and from oblique viewing angles (e.g., James and Robson, 2014).

As we find the difference in calculated resolution and subsequent grain-size measurement to be negligible between orthorectified and raw near-nadir imagery at these scales, the use of orthoimagery is not strictly necessary when using image-segmentation algorithms like *PebbleCounts* (e.g., Carbonneau et al., 2018). However, on very rough surfaces with cast-shadows from large grains, generating orthoimagery will overcome distortions present in the raw photos. Furthermore, georeferenced orthomosaics may be preferable for capturing large sites at a constant resolution that can be fed into the algorithm.

In terms of camera and photographic height (and thus resolution) considerations, one first needs to assess the minimum grain size that is desired. Following this, the resolution of the image can be determined using eq. 4.1 with some knowledge of the camera parameters (focal length, camera height, sensor size, and image size). The smallest grain b-axis needed should be 20-times this resolution. For instance, using a similar camera to the Sony $\alpha$6000 (24 megapixel, 15.6×23.5 mm sensor, 16 mm focal length), to measure all grains down to 1 cm one needs a resolution of 0.5 mm/pixel, and thus a maximum camera height of ~2 m. If finer grain sizes are desired, the user can use higher resolution imagery, but must be aware of the longer time needed for processing.

## UAV Surveying

The > 20 m flight heights typical of UAV surveys lead to cm-scale imagery with currently available 12–24 megapixel cameras, which is less appropriate for *PebbleCounts* processing, unless large (> 0.2 m) cobbles and boulders dominate the river site. Acquiring 0.5 mm/pixel imagery from a DJI Mavic drone with a 12 megapixel camera requires a very low flight height of ~1.4 m, giving a field of view of only ~1.5×2 m. This may be improved using better cameras like on the Mavic 2 Pro (20 megapixel camera), but gathering such imagery with the high overlap (~80%) required for SfM-MVS processing is still difficult, particularly given current ~20-minute flight length limitations from available batteries. Given continual technology improvements (e.g., greater battery life, more accurate geo-tags from onboard dGPS, higher megapixel cameras, and reduced motion blur), it is within reason to expect hectare to multi-hectare SfM-MVS UAV surveys at mm resolution in seamless orthomosaics along entire river reaches in the

near future. But, for the time-being, a single, non-overlapping orthoimage workflow proposed by Carbonneau et al. (2018) has high potential to achieve large-areal results. Their workflow, building on Carbonneau and Dietrich (2017), uses a number of high and oblique overlapping flights to orthorectify a lower non-overlapping flight with mm-scale acquisition, with resulting single, scaled images passed to *Basegrain*, or, alternatively, to *PebbleCounts*.

## Coverage and Processing Limits

Using handheld imagery, a survey site of 1,000–5,000 $m^2$ with ~10 GCPs measured via dGPS can be covered in 2–6 hours by one person (including GCP collection). Using a camera-on-mast setup, this time can be reduced by half, with even greater speed possible using more people and cameras (of the same sensor dimensions, focal length, and height). The potential to cover even larger sites up to or exceeding 100×100 m (1 hectare) is feasible in a day of work by two people (with one measuring the targets and both sharing the photo-taking) using the proposed method with a 16–20 mm focal length lens and a 3–5 m mast.

One limit of the scalability of the *PebbleCounts* method is processing time. The KMS *PebbleCounts* tool is recommended to be applied to maximum 1–2 $m^2$ patches, depending on the image resolution, as the manual clicking of good grains is time consuming, requiring 5–20 minutes per patch depending on patch size, image resolution, and abundance of finer grains. On the other hand, the AIF *PebbleCountsAuto* tool can theoretically be applied at larger scales. However, it is also advisable to tile data and feed it to the algorithm in maximum 1–2 $m^2$ patches for ~1 mm/pixel imagery, since the non-local means denoising can take minutes on very large images (> 2,000×2,000 pixels). Again, the use of systems with GPUs or large memory will shorten processing times and allow for larger images to be run.

In practical terms, a workflow to cover a ~2,500 $m^2$ survey site captured at 1 mm/pixel resolution and processed into a georeferenced orthomosaic would be: (1) tiling into 2 $m^2$ patches, (2) passing each patch to the AIF *PebbleCountsAuto* tool with quick manual steps of shadow-masking and sand-clicking (if sand is present), where each tile takes 1–2 minutes, (3) selecting a random subset of ~20 tiles to pass to the KMS *PebbleCounts* tool as validation and uncertainty estimation for the AIF approach. Such a workflow could be accomplished in 1–2 days of work by an experienced user, providing tens- to hundreds-of-thousands of measured grains from the survey site and a robust measurement of the full grain-size distribution. To increase processing speed, a gridded subset of tiles could also be extracted from the full survey site, with a 3–5 m step size between patches, to provide complete coverage across heterogeneous gravel-bar features, while avoiding unnecessary over-sampling and processing of every patch in the survey site.

# Conclusions

Using a k-means approach for pebble segmentation in the spectral and spatial domain combined with fast manual selection of good results, we developed a new semi-automated algorithm for grain sizing optimized for images taken over gravel-bed rivers (*PebbleCounts*). We also developed an automated algorithm that uses suspect grain filtering (*PebbleCountsAuto*), albeit with larger uncertainties in the results. The lower truncation of the methods (minimum b-axis length measurable) is limited to 20-pixels and above. These new methods were necessary to acquire grain-size distributions from dynamic high-mountain rivers with complexity from sources such as large ranges in grain

size, intra-granular heterogeneity, grain overlap, irregular shadowing, and sand patches. Similar to previous methods, *PebbleCounts* is best applied at the patch scale (~1 $m^2$), however, *PebbleCounts* provides more realistic results in complex images without any post-processing steps in ~5–10 minutes per patch, assuming ~1 mm/pixel resolution imagery. *PebbleCountsAuto* performs very well on high-quality (low-blur) imagery, though with remaining misidentification that must be approached with caution. Grain-sizing results can be upscaled to areas on the order of $10^2$–$10^4$ $m^2$ when *PebbleCounts* results are used as validation for the automated *PebbleCountsAuto* function.

# Code Availability

*PebbleCounts* is a Python based program with the code and documentation available on GitHub at: https://github.com/UP-RS-ESP/PebbleCounts (Purinton and Bookhagen, 2019a).

# Acknowledgements

Anna Rosner is thanked for assistance with fieldwork for mast surveys. Steffen Wellegehausen is thanked for aiding in the lab experiment setup. Funding was sourced from DFG funded IRTG-StRATEGy (IGK2018) and NEXUS funded through the MWFK Brandenburg, Germany, both for Bodo Bookhagen. We acknowledge the support of the Open Access Publishing Fund of the University of Potsdam. Constructive reviews from Patrice Carbonneau, Pascal Allemand, and Eric Lajeunesse improved the structure of the manuscript.

# Supporting Information

See Appendix B.

Chapter 5

# Multiband (X, C, L) radar amplitude analysis for a mixed sand- and gravel-bed river in the Eastern Central Andes[†]


## Abstract

Synthetic Aperture Radar (SAR) amplitude measurements from spaceborne sensors are sensitive to surface roughness conditions near their radar wavelength. These backscatter signals are often exploited to assess the roughness of plowed agricultural fields and water surfaces, and less so to complex, heterogeneous geological surfaces. The bedload of mixed sand- and gravel-bed rivers can be considered mixed smooth (compacted sand) and rough (gravel) surfaces. Here, we assess backscatter gradients over a large high-mountain alluvial river with aerially exposed sand and gravel bedload using X-band TerraSAR-X/TanDEM-X, C-band Sentinel-1, and L-band ALOS-2 PALSAR-2 radar scenes. In a first step, we compare backscatter response over vegetation-free endmember surfaces outside of the channel in the study catchment, and also to four representative field surveyed channel cross-sections. We then develop methods to extract smoothed backscatter gradients downstream along the channel using soil-moisture correction and kernel density estimates. In a final step, the presence of sand and gravel bars is analyzed using Fourier frequency analysis, by fitting stretched exponential and power-law models to the power spectrum. We find a large range in backscatter depending on the heterogeneity of contiguous smooth- and rough-patches of bedload material. The SAR signal responds primarily to the fraction of smooth-sand bedload, but is further modified by gravel elements. The sensitivity to gravel is more apparent in longer wavelength (L-band) radar. Because the spatial extent of smooth sand bars is typically < 50 m, only higher resolution sensors (e.g., TerraSAR-X/TanDEM-X) are useful for power spectrum analysis. Our results show the potential for mapping sand-gravel transitions and local geomorphic complexity using SAR amplitude.


## Introduction

High-mountain rivers act as the drains of orogenic belts by transporting eroded rock material from hillslopes downstream as detritus in mixed sand and gravel bedload. The characteristics of this sediment, and the geological deposits

[†] *Purinton, B. and Bookhagen, B.: Multiband (X, C, L) radar amplitude analysis for a mixed sand- and gravel-bed river in the eastern central Andes, 2019, in review.*

they form, hold keys to deciphering drivers of mountain building and erosion (e.g., Strecker et al., 2007; Attal et al., 2015; Dingle et al., 2017). Variation in bedload quantity and grain-size distributions control channel geometry over decadal to millennial timescales (e.g., Sklar et al., 2006; Pfeiffer et al., 2017). This bedload has implications for ecological management (e.g., Kondolf and Wolman, 1993) and provides essential nutrients for agriculture in floodplains. In turn, the impact of human infrastructure on rivers can lead to significant changes in sediment flux (Syvitski et al., 2005) and bedload characteristics (Grant, 2012), with far-reaching implications (Walter and Merritts, 2008).

As rivers flow down a mountain catchment and into the foreland, they pass through tectono-geomorphic zones created by geological setting, climate, vegetation, and drainage network connections (e.g., Rice and Church, 1998; Church, 2002; Attal and Lavé, 2006). These zones have differing channel morphology and cause heterogeneous mixtures of sand- and gravel-dominated bedload. We refer to this as sand-gravel patchiness (i.e., discrete patches, or bars, on the order of 1–100 $m^2$ with similar bedload characteristics). The arrangement and size of sand and gravel bars can vary significantly over short distances both along- and across-channel, making sampling characteristic grain sizes difficult via traditional measures (e.g., Wolcott and Church, 1991). Digital grain-sizing techniques for high-mountain rivers are improving (e.g., Purinton and Bookhagen, 2019a,b), but capturing channel-bedload characteristics at the scale of mountain belts remains challenging (Dingle et al., 2017).

Herein lies the utility of remote sensing with spaceborne platforms. Specifically, the response of Synthetic Aperture Radar (SAR) backscatter intensity (the amplitude of returned signals) is modulated by a combination of radar-look geometry, surface roughness, and soil-moisture content of the surface (Ulaby et al., 1982; Farr, 1993). Roughness, which dominates returns on non-vegetated surfaces (Dierking, 1999), is often defined as either the root mean square of height variations ($Hrms$) and the auto-correlation length ($L$) within a cell (e.g., Aubert et al., 2011), or, in some cases, as the power spectrum slope and offset from cm-resolution 1-D surface profiles (e.g., Weeks et al., 1996). Although neither method provides perfect characterization of scale-dependent natural roughness (Shepard et al., 2001), empirical and theoretical models between these parameters and backscatter have demonstrated positive relationships (e.g., Dierking, 1999).

One way to consider the patchy mixture of sand and gravel in a riverbed is as a mixed smooth (sand bars) and rough (gravel bars) surface. The Rayleigh criterion of electromagnetic interaction with surfaces places the smooth-rough transition at approximately $\lambda/(8 * cos(\theta_i))$, where $\lambda$ is the radar wavelength and $\theta_i$ is the local incidence angle of the wave (Farr, 1993). Where surface roughness is significantly below the radar wavelength (e.g., 3.1 cm for X-band), the radar signal reflects off of the surface away from the sensor leading to low backscatter intensity measured (e.g., on compacted sand or calm water bodies). As roughness increases to the scale of the wavelength (e.g., contiguous gravel patches), increasing radiation is scattered back towards the sensor with a nonlinear relationship up to a plateau beyond which increasing roughness leads to little change in backscattering (i.e., the surface is "saturated"; Campbell and Shepard (1996)). Peake and Oliver (1971) defined two cutoffs for radar-smooth and radar-rough surfaces as $\lambda/(25 * cos(\theta_i))$ and $\lambda/(4.4 * cos(\theta_i))$, respectively. These criteria are useful references for expected transitional behavior of the SAR signal on gravelly surfaces (McCauley et al., 1982).

Geological surface roughness has been investigated to some extent using backscatter intensities over dry desert surfaces (e.g., Evans et al., 1992; Farr and Chadwick, 1996; Ridley et al., 1996; Campbell, 2001; Weeks et al., 1996, 1997; Deroin et al., 1997; Hetz et al., 2016), however, most recent radar backscatter research has focused on agricultural settings to measure soil roughness and moisture (e.g., Baghdadi et al., 2008, 2018; Rahman et al., 2008; Srivastava et al., 2009; Aubert et al., 2011; Vreugdenhil et al., 2018). Within this, much work has gone into the theoretical and empirical modeling of radar response to rough surfaces in order to retrieve the soil-moisture signals of primary interest (e.g., Baghdadi and Zribi, 2006; Bryant et al., 2007; Verhoest et al., 2008; Gorrab et al., 2015).

Few studies have applied SAR roughness measurements to fluvial environments, and all have been limited to dry desert beds and sparse observations (McCauley et al., 1982; Baade and Schmullius, 2010; Sadeh et al., 2018). In this study, we apply radar backscatter measurements for three radar wavelengths (X-, C-, and L-band) to measure downstream changes in bedload grain size along a 120-km reach of a high-mountain sand- and gravel-bed river in the eastern central Andes. Our results demonstrate that different radar wavelengths and spatial resolutions are useful for deciphering sub-pixel roughness and measuring the sand-gravel patchiness of dynamic mountain rivers.

# Study Area and Datasets

## Río Toro

We selected the mixed sand- and gravel-bed Río Toro in Northwestern Argentina as our study site based on past fieldwork and availability of data time series (Fig. 5.1). The 120-km main-stem of the river passes through a steep environmental, topographic, and rainfall gradient from the upper reaches (~3 km elevation) bordering the hyper-arid (rainfall < 0.2 m/yr), internally drained Altiplano-Puna Plateau (Allmendinger et al., 1997; Bookhagen and Strecker, 2008), down through a zone of large Quaternary fill-terraces (Tofelde et al., 2017), then a narrow bedrock gorge (Hilley and Strecker, 2005), and out through the orographic barrier at the mountain front (rainfall > 1 m/yr), before passing through the Lerma Valley (~1 km elevation). The topographic, climatic, and environmental zones are captured in profile-view along the channel in Figure 5.2, showing rainfall, inside- and outside-channel Normalized Difference Vegetation Index (NDVI), and channel-bed slope and width. High-intensity rainfall events, particularly during the summer monsoon season (Castino et al., 2016a,b, 2017), lead to flooding and bedload transport between longer periods of low-flow and bedload aerial exposure in the braided channel networks. Vegetation in the active channel is limited because of the high bedload transport (Purinton and Bookhagen, 2018), although in the more arid, but low-slope, upstream reaches there are some scattered (~2–10 m separated), low (~0.5–1.5 m tall) bushes in the channel, which cause local increases in NDVI (Fig. 5.2B).

## SAR Datasets

In the X-band ($\lambda = 3.1$ cm), we used 12 TerraSAR-X/TanDEM-X (TSX/TDX) dual horizontal (HH) polarized stripmap scenes, with eight from descending passes and four from ascending, all processed to ~5-m ground resolution with 2×2 (range×azimuth) multilooking (Krieger et al., 2007). The C-band ($\lambda = 5.6$ cm) data are from Sentinel-1A/B (S1) dual vertical (VV) and vertical horizontal (VH) polarized sensors in stripmap mode, with 15 scenes separated into five descending and 10 ascending, processed to ~15-m ground resolution using 4×1 multi-

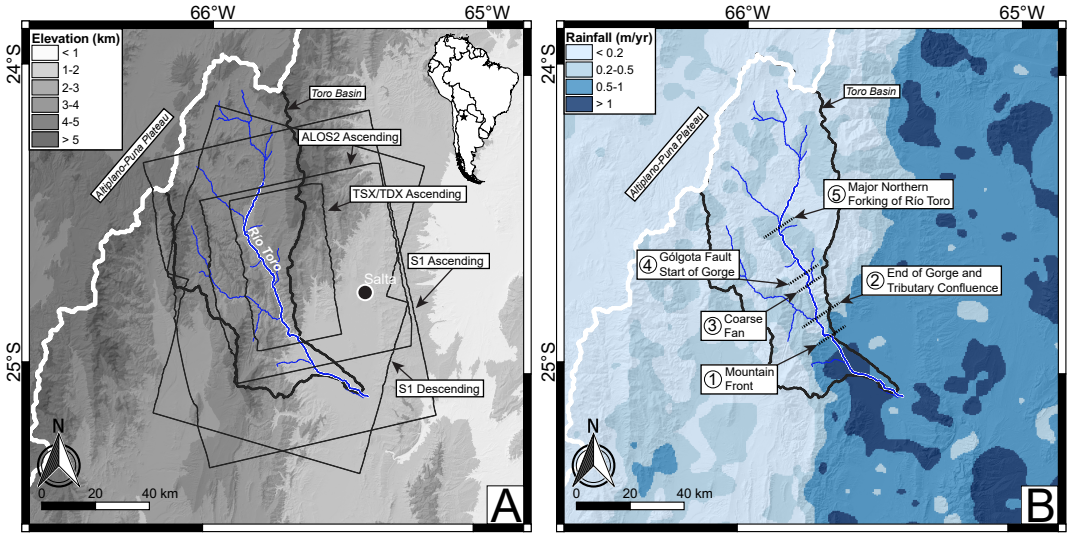


**Figure 5.1** – Overview of the study area. (A) shows the topography of the Toro Basin (thick black outline) in Northwest Argentina near Salta, bordering the internally drained Altiplano-Puna Plateau (thick white outline) of the central Andes. The Río Toro main-stem measured area (120-km length) is highlighted by the thick blue line, and representative SAR footprints used in this study are shown with thin black outlines. The elevation, hillshade, and drainage network is derived from SRTM data (Jarvis et al., 2008). (B) shows the mean annual rainfall over 12 years from the Tropical Rainfall Measurement Mission (TRMM2B31; Bookhagen and Strecker (2008)) to highlight the steep environmental gradient from the arid, sparsely vegetated upper Toro to the humid, vegetated foreland, with orographic moisture blocking indicated by the light (west) versus dark (east) color difference at the mountain front running from south to north. The locations of tectono-geomorphic transitions along the channel are indicated by dashed black lines with the numbers (1–5) used in subsequent plots.

looking (Torres et al., 2012). Finally, we have three L-band ($\lambda = 22.9$ cm) scenes from the ALOS-2 PALSAR-2 (ALOS2) HH and horizontal vertical (HV) polarized sensor, all from ascending passes, also processed to ~15-m ground resolution using $5 \times 2$ multilooking (Kankaku et al., 2013). A table with scene dates can be found in Appendix C (Table C1).

## Use of Radar Over Optical Data

Previous research has demonstrated the utility of optical data in assessing grain-size information from sub-pixel shadowing (Weeks et al., 1996; Mushkin and Gillespie, 2005, 2006). In this study we focus on radar imagery, because optical techniques are highly dependent on lighting (sun angle) and atmospheric conditions (cloudiness), and require two scenes gathered at the same time from different angles (Mushkin and Gillespie, 2005, 2006). On the other hand, radar data are capable of penetrating cloud cover and can retrieve backscatter information at any time of day and using only one scene. Furthermore, where shadowing is a correlative measure of surface roughness, radar more directly measures the relative height of the surface from reflected amplitude signals. However, radar backscatter data are influenced by several parameters, including vegetation and soil moisture (e.g., Verhoest et al., 2008; Vreugdenhil et al., 2018), thus requiring caution in analysis. The difference between optical and radar data in measuring channel-bedload characteristics is demonstrated in Figure 5.3.

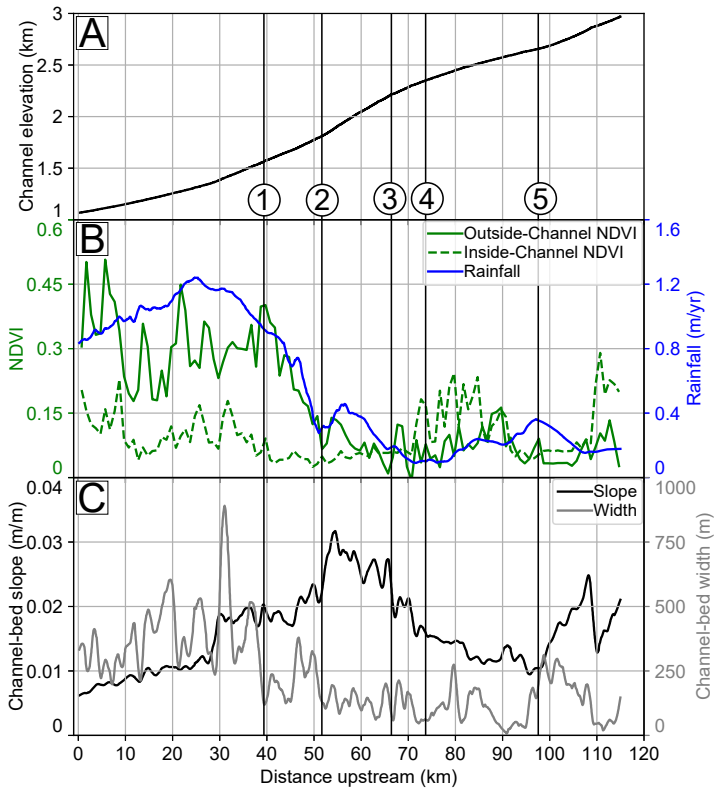


**Figure 5.2** – Topographic, climatic, and environmental characteristics of the Río Toro in profile. (A) shows the elevation profile. (B) shows mean TRMM-derived rainfall in a 5-km-wide swath along profile, and the mean inside- and outside-channel Normalized Difference Vegetation Index (NDVI) in 1-km bins along the channel. (C) shows the channel-bed slope and width, both smoothed using a Savitzky-Golay filter with a second-order polynomial and a window length of 3 km. The tectono-geomorphic transitions (1–5) correspond to Figure 5.1B. Note the steepening of the channel downstream of (4) where the Gólgota fault (Marrett et al., 1994) crosses the channel and creates a broad knickzone leading into a high-relief bedrock gorge, where many small tributaries and steep hillslopes deliver coarse material to the channel bed, such as the coarse fan at (3). Following the confluence of the large tributary at (2) and end of the gorge, the channel shallows and widens downstream of the mountain front at (1) and out into the foreland, with corresponding increases in rainfall and vegetation (mostly farmland, with intermittent dense tree stands) outside of the channel bed.

# Methods

## SAR and NDVI Data Preprocessing

All SAR data were received in L1 Single Look Complex (SLC) format and processed to square gridded pixels using the Sentinel Application Platform (SNAP) (SNAP, 2019). Two NDVI rasters were generated using GDAL (GDAL/OGR contributors, 2019) for the Toro Basin using 10-m resolution Sentinel-2 (S2) scenes ($NDVI = (band8 - band4)/(band8 + band4)$) from winter (20 July 2017) and summer (11 November 2017). Prepro-

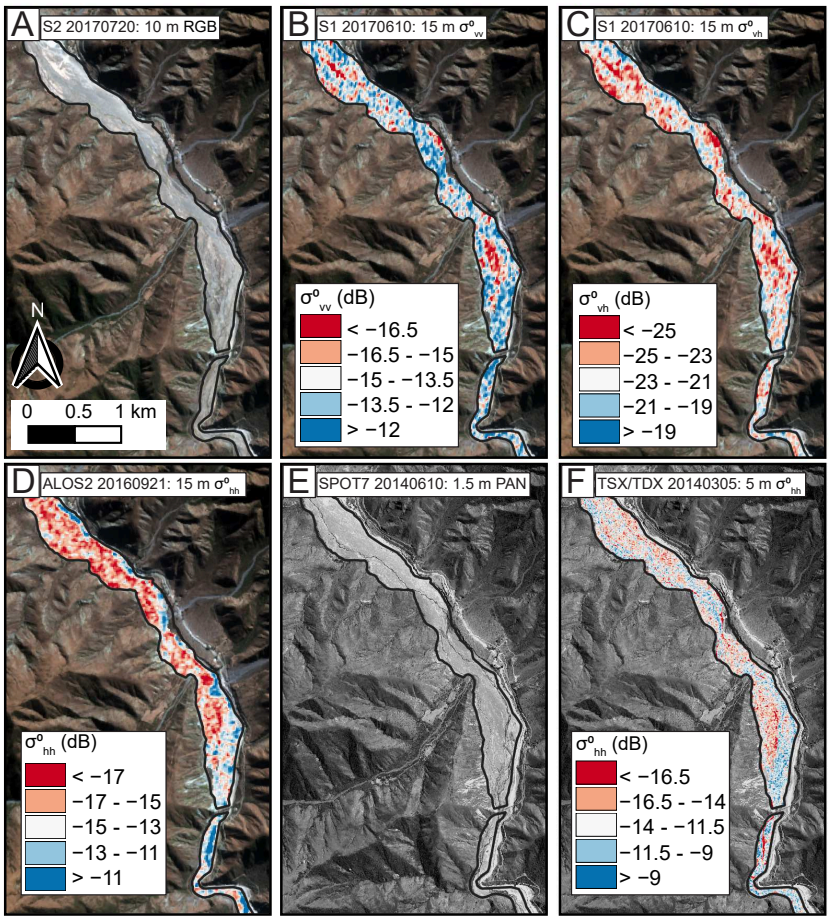


**Figure 5.3** – Comparing optical versus radar data in the Río Toro. (A) shows a Sentinel-2 (S2) RGB composite (bands 2, 3, and 4) with (B), (C), and (D) showing the same 10-m resolution optical scene overlain with 15-m resolution Backscatter from S1 $\sigma^0_{vv}$, S1 $\sigma^0_{vh}$, and ALOS2 $\sigma^0_{hh}$, respectively. (E) shows a SPOT7 panchromatic (PAN) 1.5 m image for the same area and in (F) the corresponding 5-m TSX/TDX $\sigma^0_{hh}$. The channel (hand-clicked black outline), flows from north to south. Calculation of the backscatter coefficient ($\sigma^0$) is described in the methods section. All color scales are from the $1^{st}$–$99^{th}$ percentile of backscatter values in the frame. We note some visible differences in optical reflectance that could be exploited, but point out that the information in the SAR signal is more detailed, with lower (more red) values indicating smoother surfaces and higher (more blue) values indicating rougher surfaces.

cessing of the SAR amplitude to backscatter coefficient measurements was done in the following steps, where each step is labeled TSX/TDX, S1, and/or ALOS2:

1. Thermal noise removal using a dataset provided look-up table (S1).
2. Radiometric calibration of the amplitude values to the backscatter coefficient (TSX/TDX, S1, ALOS2).
3. Debursting to combine the three image swaths (S1).
4. Application of a precise orbit file to improve the positional accuracy of the scene (S1).

5. Speckle (salt and pepper) noise filtering using a 3×3 boxcar (mean) filter (TSX/TDX, S1, ALOS2).

6. Multilooking the range and azimuth oblong pixels to approximate square pixels and further reduce noise (TSX/TDX: 2×2, S1: 4×1, ALOS2: 5×2).

7. Terrain flattening and terrain correction using the SRTM-C 1 arcsec digital elevation model (DEM) to convert the backscatter coefficient to the terrain-corrected $\gamma^0$ value (Small, 2011) (TSX/TDX, S1, ALOS2).

8. Writing a multiband GeoTiff for each scene with square pixels containing $\gamma^0$ in each available polarization (HH, HV, VV, and/or VH) and the local incidence angle from the DEM (TSX/TDX: 5 m $\gamma^0_{hh}$, S1: 15 m $\gamma^0_{vv,vh}$, ALOS2: 15 m $\gamma^0_{hh,hv}$).

Resulting backscatter and NDVI rasters were clipped to: the Río Toro hand-clicked outline; three endmember surfaces in the study area (Fig. 5.4); and a buffered region outside the channel (100-m distance from either side and 100-m wide) used to correct for soil-moisture effects. We converted the backscatter coefficient to the local incidence angle corrected $\sigma^0$ value, where $\sigma^0 = \gamma^0 * cos(\theta_i)$ (Small, 2011), expressed in logarithmic decibel (dB) format, where $dB = 10 * log_{10}(\sigma^0)$. We focus on the dual-polarized band from each sensor, but we reference the cross-polarized results and include figures in Appendix C for S1 (VH) in Figures C4 and C9 and ALOS2 (HV) in Figures C6 and C11.

# Surface Roughness Characterization

We do not provide a detailed characterization of roughness via $Hrms$, $L$, and/or high-resolution 1-D elevation profile power spectra parameters (Dierking, 1999; Shepard et al., 2001). Such efforts may be appropriate at the plot-scale in relatively homogeneous settings where roughness data are collected at or near the time of SAR measurement (e.g., Weeks et al., 1996; Baghdadi et al., 2008; Aubert et al., 2011), but our studied channel is heterogeneous and large, and our SAR scenes span from 2011–2019 (Table C1). Therefore, detailed parameterization of individual sites (i.e., a few SAR pixels) in the field is less useful than a study of the relative differences in downstream backscatter intensity based on field observations and geomorphic knowledge. We use the terms smooth and rough relatively, as they relate to the patchy arrangements of the smooth sand-surface fraction (e.g., sand bars where height variations are < 0.5 cm) and rough gravel elements (e.g., gravel bars covering a large range of grain size from < 1 cm to > 50 cm in diameter).

# Backscatter Intensity Analysis

## Backscatter Over Endmembers

We first assessed the response of the X-, C-, and L-band wavelengths over representative smooth and rough endmember surfaces in the study area. These were: homogeneously rough Quaternary terraces, a heterogeneously rough gravel pile, and a homogeneously smooth sandy outwash fan. We took the average and standard deviation of $\sigma^0$ using every pixel from every scene covering the endmember. The average and propagated standard deviation of each scene individually provided comparable results to combining all scenes.

## Backscatter Over Field Sites

To determine the expected response in the sand- and gravel-bed channel, we extracted the surface sand fraction and grain-size distribution from channel-bed photo-surveys conducted in March 2018 and 2019. This was done via image processing approaches described in Purinton and Bookhagen (2019b), referred to as *PebbleCounts*, involving k-means segmentation in the spectral and spatial domains and automatic edge detection and grain delineation. We compared these measurements with S1 data, which provided the greatest coverage over the channel (see footprints in Figure 5.1A). The short repeat pass time and continual collection of freely available S1 scenes provided data from within one week of the surveys. These S1 measurements therefore represent the $\sigma^0$ response of the gravel and sand arrangement present at time of survey collection as no bank-full floods occurred. The field sites were ~1000–4000 $m^2$ in size, allowing the extraction of 3–19 pixels at 15-m resolution for comparison of $\sigma^0$ with the sand fraction (a proxy for site smoothness) and the $95^{th}$ percentile of gravel size (a proxy for site roughness).

## Downstream Trends in Backscatter

We extracted all $\sigma^0$ pixels for each sensor and scene separately from within the 120-km hand-clicked channel reach to evaluate trends in the signal. All pixels with seasonal NDVI values > 0.15 were masked to remove all but the sparsest vegetation. Furthermore, we masked any pixels that had $\theta_i$ values < 25° or > 50°, since very high or low angle backscatter returns may be influenced more by look geometry than surface roughness (e.g., Farr, 1993; Baghdadi et al., 2008; Aubert et al., 2011). The remaining $\sigma^0$ values were plotted versus upstream distance.

Because of the high data density and variability, we used a 2-D Gaussian Kernel Density Estimate (KDE) to identify the dense and sparse zones of measurement along the channel. Considering the KDE as a 3-D surface with the X and Y horizontal dimensions as distance upstream and $\sigma^0$, respectively, and the Z vertical dimension as the density, we can connect the line of maximum height (density) to extract a trendline that is independent of bin size. Remaining spikiness in the KDE trendline was smoothed using a Savitzky-Golay filter with a second-order polynomial and a window size of 201 points, where each point is one pixel, resulting in a ~3-km and ~1-km smoothing window for the 15-m and 5-m data, respectively.

We also sought to correct for effects of soil moisture on the smoothed inside-channel $\sigma^0$ trendline from each sensor and scene. Soil moisture is known to cause an increase in backscatter intensity (e.g., Verhoest et al., 2008; Gorrab et al., 2015; Bousbih et al., 2017) up to very wet conditions (> 30% moisture), where the relationship has been shown to decrease (Baghdadi et al., 2008). Although moisture in the channel bedload is very low in the arid upstream reaches of the channel, this increases downstream. We removed this signal using the outside-channel buffer zone clipped from each scene. We follow the same steps of NDVI masking, high and low $\theta_i$ removal, plotting versus upstream distance, and smoothed KDE trendline extraction. We then differenced the trendlines (inside-channel−outside-channel), to provide corrected channel trendlines dominated by the bedload roughness response.

In the outside-channel case, we used a more relaxed NDVI threshold of 0.35, allowing sparse vegetation like shrubs, grass, and some crops to remain. This insures enough pixels for the analysis outside of the channel where there is more vegetation, particularly downstream of 50 km where mean outside-channel NDVI is > 0.15 (Fig. 5.2B). We expect the radar response to the remaining sparse vegetation to be minimal (or non-existent upstream of 50 km where NDVI is < 0.15). This expectation is especially warranted in the dual-polarized configuration where

volume scattering effects of vegetation are weaker (e.g., Vreugdenhil et al., 2018). Also, we expect limited response of the L-band to these remaining low-vegetation pixels due to penetration by the longer wavelength.

## Sand-Gravel Patchiness From Backscatter Power Spectrum

To analyze the patchiness of sand and gravel bedload we remain in X-Y plan-view space to maintain the relation of each pixel to one another. For this, we use the 2-D Discrete Fourier Transfrom (DFT) to examine the frequency distribution of backscatter. This technique has been widely applied in the geosciences, for example for measuring landscape scaling relationships (e.g., Perron et al., 2008) and DEM artifact identification (e.g., Arrell et al., 2008; Purinton and Bookhagen, 2017). The 2-D DFT transforms a grid of evenly spaced values ($\sigma^0$ in our case) from the spatial to the frequency domain. This provides information on the amplitude and periodicity of the values. The power spectrum, magnitude squared of the DFT ($|DFT|^2$), is a measure of the variance of $\sigma^0$ with the units of amplitude squared ($dB^2$), and can be plotted against frequency, or wavelength ($frequency^{-1}$), to evaluate trends in the spatial signal (e.g., Booth et al., 2009).

Since the 2-D DFT requires void-free grids, we preprocessed our channel data using standard methods (Perron et al., 2008; Jacobs et al., 2017). We began by taking 1-km channel clips, insuring enough pixels for the analysis while avoiding frequent channel bends, which lead to non-rectangular clips. Then, small holes (e.g., where vegetation was masked) were filled by the Fast Marching Method interpolator (Telea, 2004) implemented in the openCV library for Python (Bradski, 2000). Filled hole sizes were limited to 20 and 10 pixels for the 5- and 15-m scenes, respectively. The channel clip was then rotated by 1° increments using bilinear resampling until the void-space surrounding the semi-rectangular data was minimized, and the clip was cropped to the minimum bounding rectangle. If the remaining percentage of voids was > 40% or the minimum spatial dimension was < 8 pixels, then the channel clip was not considered. For the remaining 1-km clips, voids (holes > 10 or 20 pixels and the remaining region surrounding the rotated channel) were filled with the median $\sigma^0$ value to generate a void-free rectangular grid. A least-squares plane fit was used to detrend the grid and a weighted cosine Hanning window was used to taper the edges to reduce spectral leakage (Perron et al., 2008).

Following preprocessing, the grid was passed to the 2-D DFT and the power spectrum was extracted. The frequency of each pixel in the 2-D power spectrum was calculated based on grid spacing and the flattened 1-D array of power spectrum was plotted against the frequency in log-log space. This plot demonstrates the decay in variance from long-wavelength (low-frequency) features to short-wavelength (high-frequency) features, or the spatial variability of large and small patches of similar $\sigma^0$ values (e.g., sand or gravel bars). We then fit functional models to the power spectrum to describe the rate of decay (Clauset et al., 2009).

First, we stacked all coincident 1-km channel clip power spectra to generate large point clouds for robust fitting. Using each scene and 1-km channel clip independently led to large scatter in the data collected over 9 years (Table C1). Stacked clips ignore inter-scene variability, but allow us to confidently assess integrated trends in sand-gravel patchiness. Initial testing demonstrated insufficient resolution (15 m) for the S1 and ALOS2 scenes, as much of the channel upstream of the mountain front had widths < 120 m (8 pixels) (Fig. 5.2C). Therefore, we relied only on the TSX/TDX 5-m scenes.

For our stacked TSX/TDX data, the majority of variance decay fell in the 10–50-m (2–10-pixel) wavelength range, which accounted for ~75% of all power spectrum measurements per-clip. Based on this, we log-binned the 10–50-m frequencies at their median power spectrum value for fitting to reduce noise in the raw signal. On this log-binned plot, we first fit a power-law function of the form $f(x) = x^{\alpha}$. The exponent ($\alpha$), or decay constant, is the power-law slope, which defines the decay in variance of the power spectrum over increasing frequencies. This function was fit with the minimum at each log-spaced frequency bin ($xmin$), up to the third-to-last bin to insure a robust fit across at least one magnitude of the power spectrum. From the ensemble of fits, the minimum two-sample Kolmogorov-Smirnov (KS) test statistic was used to select the best fit, or the $xmin$ at the statistically chosen start of power-law behavior (Clauset et al., 2009). We found an $xmin$ range of ~16–23 m across 59 stacked 1-km channel clips fit (see Figure C14). Based on this tight range, in a second round of fitting we chose to fix the $xmin$ at the median value of 18 m, and always fit the power-law over the 10–18-m binned range. Fixing $xmin$ allowed comparison of the $\alpha$ slope in the stacked 1-km channel clips without effects of a variable minimum frequency.

Fitting results indicated a second segment in the power spectrum at a longer wavelength (lower frequency) in the 18–50 m range. This region was not well-described by a power-law, and instead followed a stretched exponential function of the form $f(x) = e^{x\beta}$. The exponent ($\beta$) is the curvature of the exponential, which defines the rate of variance decay up to the transition to power-law behavior. Long-wavelength exponential fitting followed the KS-test statistic best-fit method (Clauset et al., 2009), but here we set the minimum wavelength ($xmax$) at the 18-m frequency bin and allowed the $xmin$ to vary, since we are most interested in the fit area near the power-law transition, and less concerned with differences in minimum frequency affecting $\beta$ comparisons.

From these two fits, we extracted short-wavelength (power-law, $\alpha$) and long-wavelength (exponential, $\beta$) exponents, along with their 1-standard deviation uncertainties. These decay constants and errors can then be plotted against upstream distance for every stacked 1-km channel clip. More negative values indicate steeper slopes or higher curvature, faster decay, and thus more sand-gravel patchiness (heterogeneity, indicating greater contiguous sand bars), and less negative values indicate shallower slopes or lower curvature, slower decay, and thus less sand-gravel patchiness (homogeneity, indicating more mixed sand and gravel bars).

# Results

## Endmembers

The representative endmember clips from the Toro Basin and the resulting average and standard deviation $\sigma^0$ dual-polarized values are shown in Figure 5.4. Given the ~35° incidence angle of the SAR scenes (Table C1), the Peake and Oliver (1971) smooth thresholds occurs at approximately 0.15, 0.29, and 1.08 cm, and the rough thresholds at 0.87, 1.65, and 6.13 cm for the X-band TSX/TDX, C-band S1, and L-band ALOS2 scenes, respectively. Where the average height of roughness elements is below the smooth threshold, we expect low $\sigma^0$ as the signal is reflected away from the sensor, and where the surface is above the rough threshold we expect high $\sigma^0$. The range between these transitional values indicates the range of expected responses, particularly on mixed smooth- and rough-surfaces.

As expected from wavelength considerations, the ALOS2 data always have a lower mean backscatter signal, since the surfaces appear more smooth. It follows that the TSX/TDX data has the highest mean values. We also note

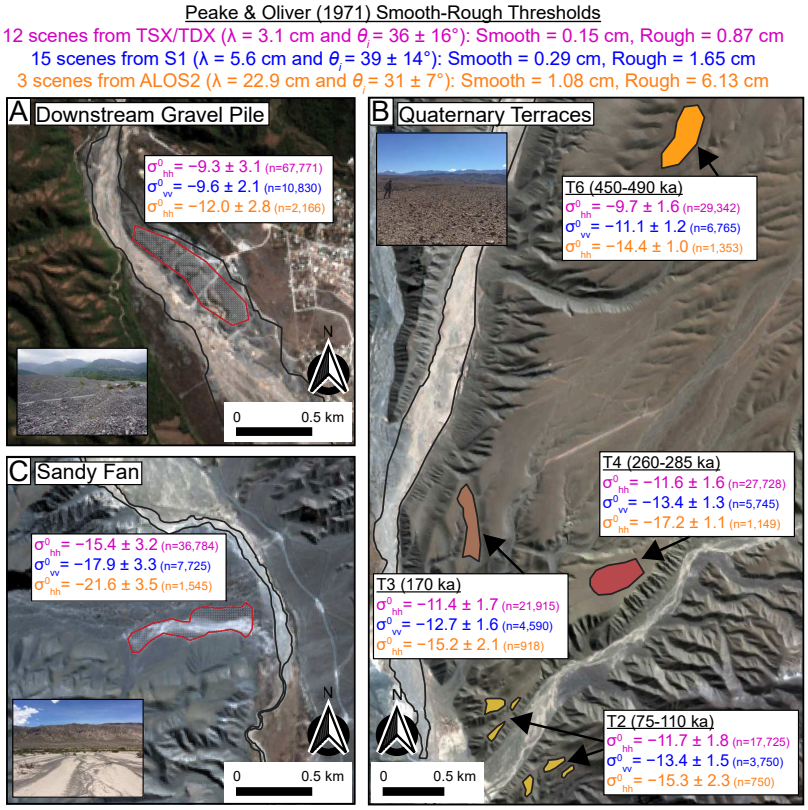


**Figure 5.4** – Differences in X-, C-, and L-band $\sigma^0$ (mean $\pm$ standard deviation in dB) from three endmembers. The locations are indicated in Figure 5.5A. (A) shows a large man-made gravel pile associated with in-channel gravel extraction, (B) shows the Quaternary terraces with dates from Tofelde et al. (2017), and (C) shows a sandy outwash fan associated with a small debris-flow prone sub-catchment. The color scheme (pink for TSX/TDX, blue for S1, and orange for ALOS2) is used consistently in the chapter and Appendix C. Note the Peake and Oliver (1971) smooth-rough thresholds listed above the figure. The river (black outline) always flows approximately north to south. Inset pictures taken from each area during fieldwork. Basemap imagery is the S2 scene used in Figure 5.3.

that the sandy fan in Figure 5.4C has the lowest mean values, as the surface is very smooth (see inset field photo). Importantly, we note the tighter $\sigma^0$ (lower standard deviation) over homogeneously rough surfaces, like the terraces (maximum 2.3 dB for ALOS2 on the youngest terrace T2, but typically < 2 dB), versus for more heterogeneously rough surfaces, like the gravel pile (maximum 3.1 dB for TSX/TDX, and all > 2 dB). The range in standard deviation from the smooth sand surface is smaller between sensors (3.2–3.5 dB) than for either the gravel pile or terraces.

# Field Sites

Moving to the channel, we turn to our field survey sites in Figure 5.5. There is a clear trend of decreasing $\sigma^0$ with increasing sand fraction, indicating the importance of these smooth sand bars in dictating the return intensity. On the other hand, the gravel size in the $95^{th}$ percentile exerts less influence on the returns with sites B, C, and D having similar gravel sizes, but a ~6 dB range in $\sigma^0$ due to differences in the sand fraction. Since the *PebbleCounts* (Purinton

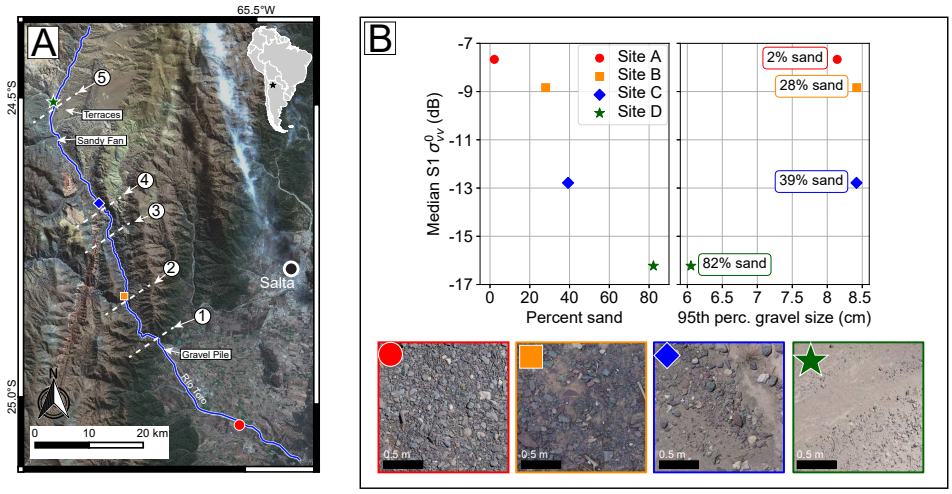


**Figure 5.5** – The S1 backscatter response over field-surveyed channel cross-sections. (A) shows the S2 RGB composite from Figure 5.3 and the measured Río Toro. The locations of tectono-geomorphic dividers are indicated by dashed white lines numbered (1–5) (Fig. 5.1B). The colored symbols correspond to the location of the field sites in (B). (B) shows the field survey sand fraction (left plot) and $95^{th}$ percentile gravel size (right plot), with a representative 2×2 m clip from each survey shown in corresponding colors and symbols below. Note the increasing sand fraction from downstream (site A) to upstream (site D). The survey sizes were 2400, 1700, 1000, and 4000 $m^2$, the number of S1 pixels were 11, 8, 3, and 19, and the number of individual gravels measured were 45,200, 15,800, 10,000, and 10,500, for sites A, B, C, and D, respectively.

and Bookhagen, 2019b) algorithms have a lower detection limit of 20-pixels in intermediate grain-size axis length, and the photo-surveys were captured at ~1 mm/pixel resolution, we are limited to measuring the gravel fraction > 2 cm. For reference, the median size in this subset of gravels was very similar at 3.1, 3.0, 3.0, and 2.7 cm for sites A, B, C, and D, respectively. Despite an expected 1.65 cm roughness-saturation transition for the S1 data (Fig. 5.4), the full range in response covers ~10 dB. From this result, we expect the backscatter response to primarily reflect the sand-bar contribution at each pixel, but to be further modulated by the presence of gravel-dominated bars, and by the size of those gravels, particularly in the case of L-band, which has a higher threshold for radar-roughness saturation of 6.13 cm.

## Downstream Trends

With a detailed understanding of expected $\sigma^0$ response over endmembers and four field sites distributed along the channel, we now examine the continuous signal along the entire channel reach. Figure 5.6 contains an example analysis from one S1 scene as discussed in the methods. (Figure C1 in Appendix C contains examples using TSX/TDX and ALOS2 scenes.) There is an increase in NDVI values within the channel downstream of the mountain front at (1) and in some low-slope reaches upstream of the knickzone at (4), with the mean 1-km-binned NDVI (Fig. 5.2B) sometimes exceeding the 0.15 channel threshold for $\sigma^0$ masking. However, many pixels remain as indicated by the raw point cloud extracted in Figure 5.6B, where we see the spikiness of the trend when using a 500-m binning approach. This spikiness is removed and we are able to extract a smooth trendline from the profile using the KDE in Figure 5.6C. The sparse measurements between (4) and (5) are indicated by the counts per 500-m bin in Figure

5.6D. Here, in the region of the Quaternary terraces, channel slopes are low and the river narrows (Fig. 5.2) and meanders at times in a wide floodplain.

KDE trendline extraction was repeated for every scene, sensor, and polarization and an additional trendline was extracted from the outside-channel buffer zone to correct for expected soil-moisture effects (see Appendix C Figures C7–C6). The plots in Figure 5.7 were generated by integrating all corrected dual-polarized trendlines from every scene for each sensor (see Appendix C Figures C7–C11). These plots were also separated by season (wet or dry) in Figures C12 and C13, and we note no difference in trend, nor is any difference noted when separating the scenes into ascending or descending passes.

These corrected trends indicate similar response of the TSX/TDX and S1 data with increasing $\sigma^0$ corresponding to decreasing sand fraction from upstream to downstream (Fig. 5.5). On the other hand, the ALOS2 L-band data show more dynamic interactions through a combination of sand fraction and gravel size, with both increasing and decreasing trends captured in local channel segments corresponding to tectono-geomorphic transitions, like the confluence of a very coarse fan at (3) delivering debris flow material to the channel. Cross-polarized results are shown in Figures C9 and C11 for S1 and ALOS2, respectively. These results are detailed more in the discussion section.

## Sand-Gravel Patchiness

The high variability demonstrated with the spiky median line and $1^{st}$–$99^{th}$ percentile envelope in Figure 5.6 indicates local geomorphic complexity of the channel bedload beyond the downstream trends. A quantitative analysis of sand-gravel patchiness is desirable, although certain localized spikes can be connected to notable features in the field; for instance, the large positive excursions between 25- and 40-km upstream in Figure 5.6B relate to the man-made gravel piles in this reach noted in Purinton and Bookhagen (2018) and shown in Figure 5.4A.

Figure 5.8 contains an example of our 2-D DFT frequency analysis for the TSX/TDX data. Combined with stacking, the high spatial resolution of the TSX/TDX data allowed fitting of power-law and exponential functions to 59 and 53 stacked 1-km clips, respectively, out of a possible 116 clips with TSX/TDX coverage. The remaining clips were either too narrow (< 8 pixels, or 40 m), contained too many void-spaces prior to median filling (> 40%), or produced poor fits ($r^2 < 0.85$, $\alpha$ fit 1-standard deviation uncertainty > 0.5, or $\beta$ fit 1-standard deviation uncertainty > 10). Although only a small range of frequency bins were fit by the short-wavelength power-law with fixed $xmin$ at 18 m (see Figure C14), this range (10–18 m) accounted for ~74% of the unbinned 10–50-m power spectrum values. The exponential fits with variable $xmin$ between 30–50 m, and fixed $xmax$ at 18 m, covered ~24% of remaining values.

From the stacked 1-km TSX/TDX channel clips, the two decay constants ($\alpha$ and $\beta$) are plotted upstream in Figure 5.9. Since this X-band sensor saturates at roughness values < 1 cm (Fig. 5.4), sand-gravel patchiness differences are chiefly related to sand bars in each 5×5-m (25-$m^2$) pixel. A more negative exponent corresponds to more sand-gravel patchiness (heterogeneity, or less contiguous sand bars), whereas, a less negative exponent corresponds to less sand-gravel patchiness (homogeneity, or more contiguous sand bars). We interpret these downstream to upstream results in the proceeding discussion.

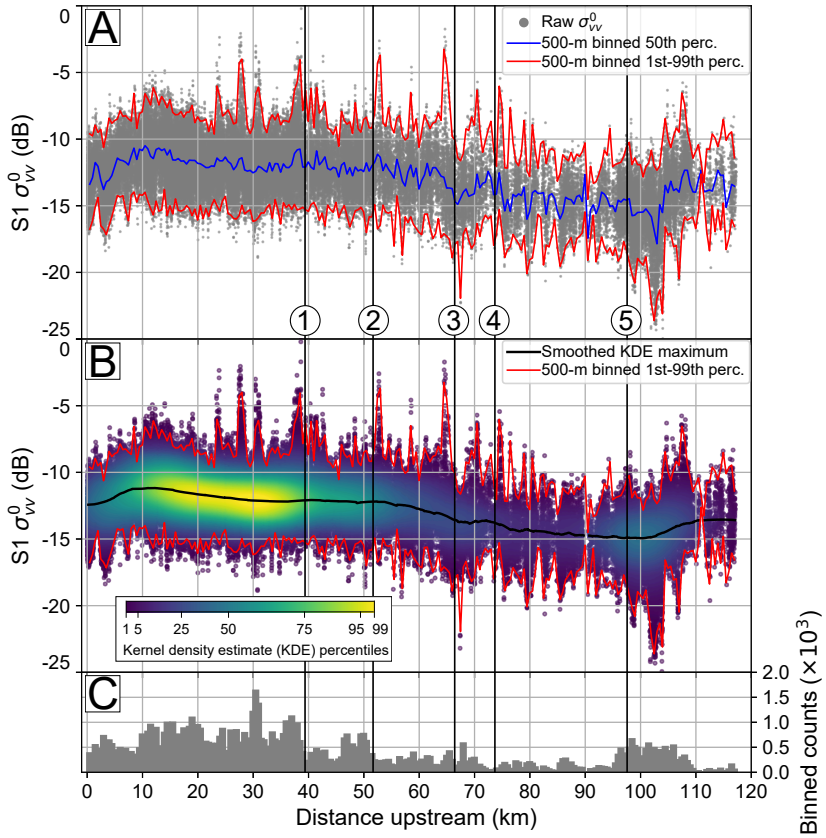


**Figure 5.6** – Example point cloud and trendline extraction for one S1 scene (20180901). The tectono-geomorphic transitions (1–5) correspond to Figure 5.1B. (A) is the raw point cloud (masked for NDVI > 0.15), with the $50^{th}$ and $1^{st}$–$99^{th}$ percentile lines (binned in 500-m steps) plotted over it. To remove the spikiness and get an average trendline we go to (B), where the median line is replaced with the smoothed 2D KDE maximum. (C) shows the 15-m SAR pixel counts per 500-m bin. See Figure C1 in Appendix C for examples using TSX/TDX and ALOS2 scenes.

# Discussion

Our results demonstrate a novel use of SAR amplitude measurements in a dynamic, high-mountain fluvial environment. The use of multiple wavelengths and resolutions show important differences and limitations of the sensors for channel-bedload measurement applications. In this discussion, we describe these differences in relation to other studies of surface roughness using SAR, point out geomorphic trends in our data, and indicate the potential uses of the method in braided alluvial rivers. We emphasize that $\sigma^0$ values are all log-scaled dB units, meaning that even small differences (~1 dB) are significant.

## Range in Backscatter for Endmembers and Field Sites

The endmember clips in Figure 5.4 show a range in mean X-, C-, and L-band $\sigma^0$ of 6.1, 8.3, and 9.6 dB, respectively. Baghdadi et al. (2008) found a similar range of X-, C-, and L-band of 4–5.5, 4, and 8 dB, respectively, on plowed agricultural fields with soil $Hrms$ values of ~0.5–3.5 cm. These relationships between $Hrms$ and $\sigma^0$ are logarith-

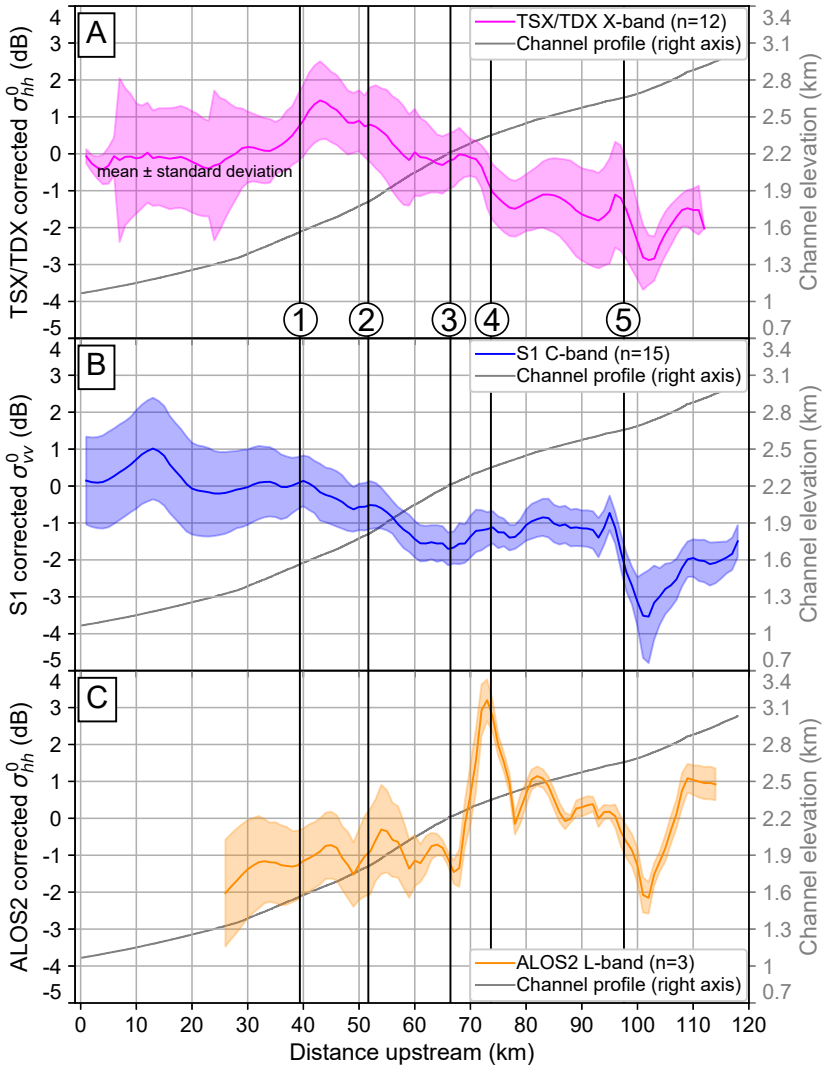


**Figure 5.7** – Smoothed KDE trendlines for each sensor from the Río Toro. To provide spatial context, each plot includes the channel elevation profile (gray line) on the right axis (see Figure 5.2), where the tectono-geomorphic transitions (1–5) correspond to Figure 5.1B. The generation of these trendlines via differencing with the outside-channel buffer is shown for the 12 TSX/TDX scenes (A) in Figure C7, for the 15 S1 scenes (B) in Figure C8, and for the 3 ALOS2 scenes (C) in Figure C10.

mic and typically reach a rapid plateau at $Hrms$ values of ~1, ~1.5, and ~2 cm for X-, C-, and L-band, respectively, beyond which the soil surface is radar saturated and increases in roughness have little effect on $\sigma^0$ (Baghdadi et al., 2008, 2018). The exact range of $\sigma^0$ and the plateau (saturation) values have varied slightly in different agricultural studies (e.g., Aubert et al., 2011; Gorrab et al., 2015; Bousbih et al., 2017), but are in general agreement.

Radar roughness studies from plowed agricultural soils with mixed sand, silt, and clay content are less applicable to our fluvial setting with mixed loose sand and gravels, where little quantitative research on backscatter response has been done. Much non-agricultural research has focused on morphological differences of dry desert alluvial fans, which show backscatter differences as a smooth desert pavement forms over millenial time-scales (e.g., Farr

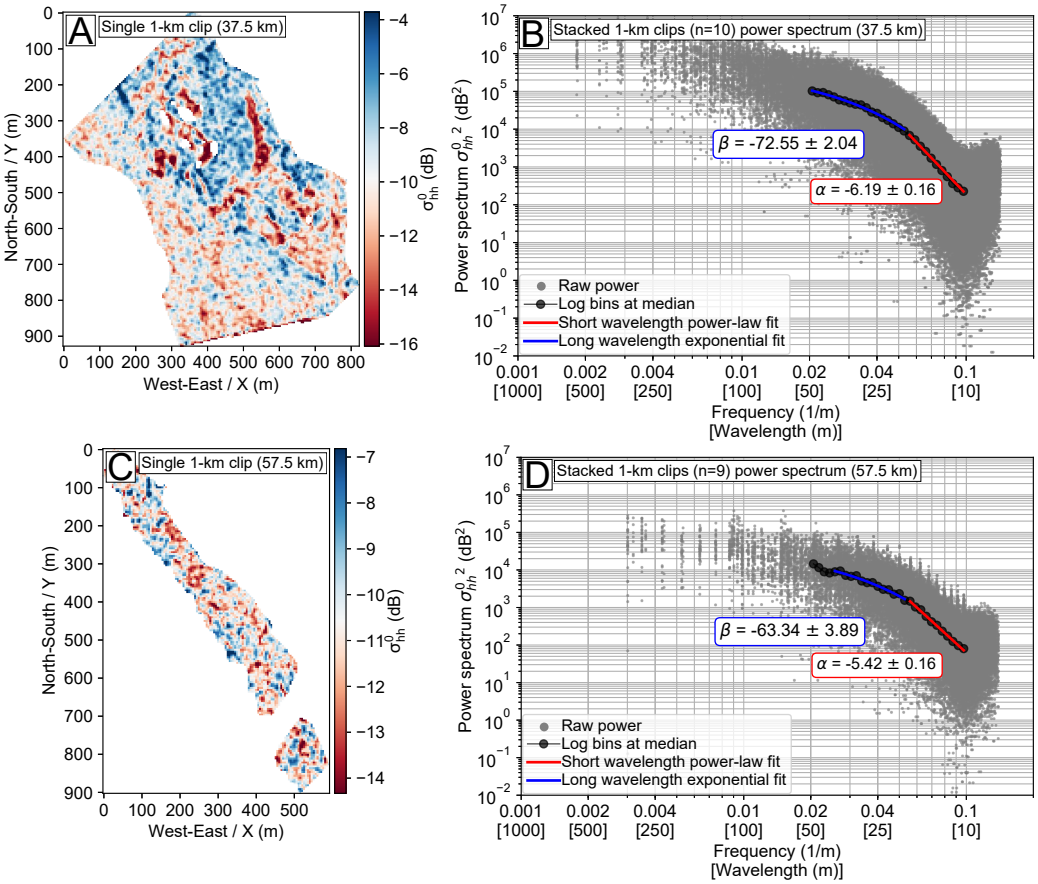


**Figure 5.8** – Example 1-km channel clips and 2-D DFT analysis from one TSX/TDX 5-m resolution scene (20150928) from 37.5 km (A) and 57.5 km (C) upstream, with $\sigma^0_{hh}$ color scale going from the $1^{st}$–$99^{th}$ percentile of all values in the clip. Each clip undergoes preprocessing to create a void-free rectangular grid (see methods section). The power spectrum in (B) shows all power-frequency raw values from 10 stacked scenes for the 1-km channel clip shown in (A), and the spectrum in (D) shows the same but for only nine available scenes covering the clip in (C). The stacking leads to a dense point cloud that can be binned and fit by one power-law (short wavelength, $\alpha$) with fixed $xmin$ at 18 m, and one stretched exponential (long wavelength, $\beta$) with a KS-test best-fit $xmin$, and $xmax$ fixed at 18 m. Note the more negative values for $\alpha$ and $\beta$ in (B) from the downstream clip (A), indicating greater sand-gravel patchiness (heterogeneity, or less sand-bar contribution).

and Chadwick, 1996; Kierein-Young, 1997; Hetz et al., 2016). Such differences in $\sigma^0$ from the different Quaternary terraces in Figure 5.4B are likely related to terrace age (Tofelde et al., 2017) and morphological differences. The tighter spread (lower standard deviation) on these terraces demonstrates the effect of a spatially homogeneous roughness, with similar rock sizes and rock spacing over a large area (see inset photo in Figure 5.4B).

Deroin et al. (1997) measured properties of loose rocks in desert test plots and found log scaling over a ~16-dB range in C-band $\sigma^0$, going from silt- to rock-dominated surfaces with a maximum rock-height range of ~0.5–15 cm (with a plateau in $\sigma^0$ beginning at ~4 cm). These ranges are greater than our 8.3-dB range from S1 C-band, and indicate the ability of extreme roughness elements on a geological surface to modulate backscatter intensity. To our knowledge, the only study of SAR bed-roughness in a hydrological sense is from Sadeh et al. (2018). Here,

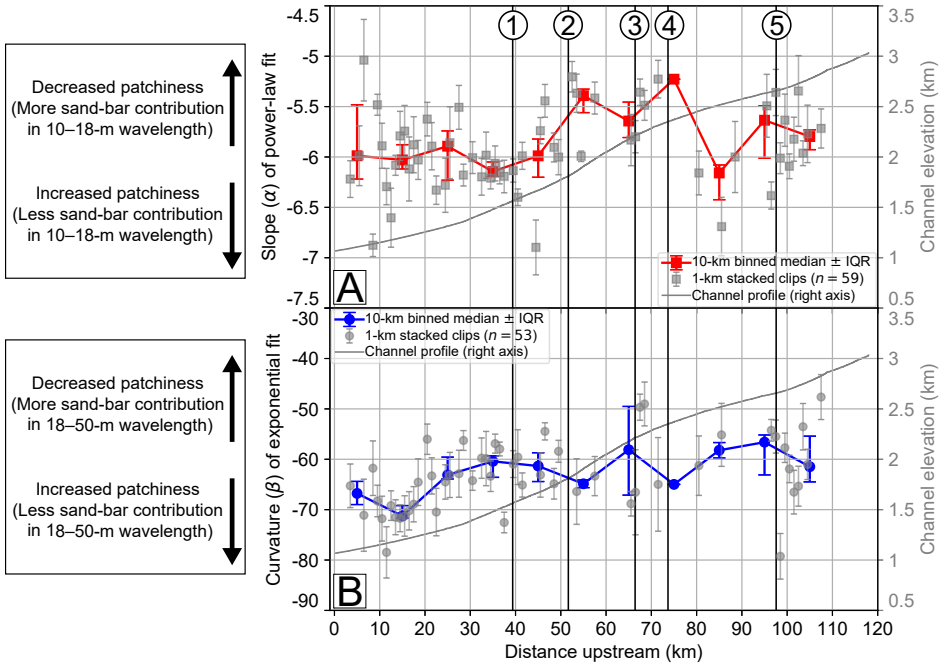


**Figure 5.9** – Downstream trends in (A) $\alpha$ and (B) $\beta$ of the power-law and stretched exponential fits, respectively, for the stacked 1-km channel clips from the 5-m TSX/TDX scenes. To provide spatial context, each plot includes the channel elevation profile (gray line) on the right axis (see Figure 5.2), where the tectono-geomorphic transitions (1–5) correspond to Figure 5.1B. Each stacked clip is shown with gray symbols and 1-standard deviation error bars. To highlight the overall trends, these exponents are aggregated in 10-km bins at the median value with per-bin Interquartile Range (IQR) error bars. More negative values correspond to less sand-bar contribution (i.e., more patchiness due to mixed gravel bars), in the 10–18-m or 18–50-m wavelength range for $\alpha$ and $\beta$, respectively.

Manning's $n$ (the hydrological roughness of a surface) was shown to correlate well with X-band backscatter. They found a ~4-dB range over rock types from desert pavement to stony limestone and a ~1.5-dB range over $Hrms$ heights from 0.7–1.7 cm.

From our channel-bed field sites (Fig. 5.5), we observe a large ~10-dB range in $\sigma^0$ from S1 C-band. The results from the channel do not indicate a clear plateau in $\sigma^0$ (e.g., Aubert et al., 2011), since the channel bed is patchy at a sub-pixel scale with areas of more or less contiguous sand bars (where height variations are typically < 0.5 cm). Each 5×5-m (25-m$^2$) or 15×15-m (225-m$^2$) pixel SAR measurement can include a mixture of gravel and sand bars that lead to a large range of possible $\sigma^0$ values, even for the C-band data covering the field sites, which should saturate at average roughness values near 1.65 cm (Fig. 5.4). Of key importance here is that differences in $\sigma^0$ are primarily related to the sand fraction (sand bars), with only minor modulation by extreme roughness elements (Deroin et al., 1997), at least in the case of C-band radar.

## Backscatter Trends and Geomorphic Implications

Our assessment of expected (from previous work) and observed (from endmember clips and field sites) backscatter response allows us to interpret trends downstream along the Río Toro. Field data on sand fraction in Figure 5.5B indicates an approximate upstream sand-dominated (> 50%) to downstream gravel-dominated transition between

the major northern river fork at (5) and start of the steep bedrock gorge at (4), between approximately 75- and 100-km upstream in the trendlines shown in Figure 5.7.

Upstream of 100 km, all three radar bands show a decrease then increase in $\sigma^0$ with a magnitude of ~1, ~1.5, and ~3 dB for the X-, C-, and L-band, respectively. This is a channel reach with many smooth sand bars, so the radar response is similar, though with different magnitude. As we move downstream of the northern fork at (5) and the gravel fraction increases, we begin to see differences between wavelengths. For the X-band TSX/TDX trendline in Figure 5.7A, we note a larger standard deviation envelope, given the high variability in these scenes captured over 5 years (Table C1). In general, TSX/TDX $\sigma^0$ increases steadily over a magnitude of ~3.5 dB from 100- to 40-km upstream at (1) (the mountain front), at which point the signal drops by ~1.5 dB and then flattens. The C-band S1 trendline in Figure 5.7B shows a somewhat similar trend with a sudden increase of ~3 dB when crossing the 100-km upstream fork at (5), then flattening and slightly decreasing by ~0.5 dB until the entrance of the coarse debris flow fan at (3), at which point the signal steadily increases by ~2 dB to the mountain front before flattening. Similar trends in X-band (3.1 cm) and C-band (5.6 cm) indicate a response dominated by sand fraction in the channel, because these shorter wavelengths saturate at roughness thresholds < 2 cm (Fig. 5.4).

The L-band ALOS2 trendline in Figure 5.7C has a very different and much more dynamic response, since this long wavelength (22.9 cm) SAR sensor — with a high roughness saturation threshold of 6.13 cm (Fig. 5.4) — reflects changes in gravel roughness elements in addition to the sand fraction. The L-band increases by ~5 dB (to its peak) from the tributary fork at (5) to the start of the gorge at (4). The signal then drops swiftly by ~4.5 dB over only ~10-km distance to the coarse fan at (3). Following this, there are a number of local ±0.5–1-dB fluctuations as minor tributaries and hillslope processes (e.g., landslides) deliver sediment packages of diverse grain size to the channel bed, which create fine and coarse gravel bars (e.g., Attal and Lavé, 2006). Despite fluctuations, the decreasing size and abundance of large boulders (> 25 cm) observed in the field (i.e., extreme roughness elements; Deroin et al. (1997)) leads to an overall ~1 dB decrease from 65-km to the limit of ALOS2 coverage (Fig. 5.1A) at 25-km upstream.

These trends are all corrected for soil moisture (Figure C7–C11), but the water surface in the channel may still modulate the roughness signal (e.g., Hwang and Fois, 2015; Uddin et al., 2019). Upstream of the mountain front the width of the active channel is limited to ~2–4 m, so there is limited effect of the water surface in the typically > 100-m wide channel bed. Downstream of the mountain front, the active channel can locally exceed ~10 m, particularly in the wetter summer months. This water is turbid with high suspended sediment loads and a rough surface that may contribute to some increases in backscatter. However, in this region the channel bed is typically > 200-m wide, in places exceeding 800 m (Fig. 5.2C), and thus these water pixels have a limited overall effect on the integrated signal of many pixels over months to years of scene collection, during which time the active channel location shifts in the braided bed. Such shifting of the active channel and rearrangement of the gravel and sand bars is likely reflected by the larger error envelopes near and downstream of the mountain front (40 km) in Figure 5.7.

As noted, the assessment of sand-gravel patchiness via power-law and exponential decay constants from backscatter power spectra in channel clips was only possible using our 5-m TSX/TDX data. This X-band radar primarily reflects the contribution of smooth sand bars, as even small gravels will rapidly saturate the signal. A large spread from the individual channel clip decay constants (Fig. 5.9) is expected given the dynamic channel, where sand bars shift

regularly between seasons and over years, even under low-flow conditions where larger gravels are not transported. Integrating 5 years of TSX/TDX scenes, including noise in the individual measurements, further contributes to the range and fitting uncertainty of $\alpha$ and $\beta$.

For the short-wavelength (10–18 m) power-law fits (Fig. 5.9A), $\alpha$ decreases from the upstream reach to the start of the gorge (patchiness increases), meaning more sand bars are present. In the steepened (due to the fault) and narrow gorge from 50- to 75-km upstream, the channel bed becomes less patchy, meaning the sand and gravel is present in more mixed bars. Downstream of the gorge, patchiness increases to the mountain front, so sand bars are again formed in the 10–18-m wavelength range in the wider, lower-slope channel bed (Fig. 5.2C). Notably, $\alpha$ clusters tightly in the 30–40-km reach, where anthropogenic modification (gravel mining; see Purinton and Bookhagen (2018)) and high bedload transport creates constant patchiness, with relatively lower sand-bar contribution. The sand-gravel patchiness scatters downstream of 30 km, where there is a diverse mix of sand and gravel bars in the low-slope, wide-channel foreland (Fig. 5.2C).

For the long-wavelength (18–50 m) stretched exponential fits (Fig. 5.9B), $\beta$ has an overall increasing downstream trend in patchiness. This indicates that in this longer-wavelength range, the presence of contiguous sand bars decreases downstream as the increased channel activity (higher flow) regularly rearranges the bedload and mixes the gravel elements at this scale. Also, there is generally a lower spread (smaller IQR) of the $\beta$ exponent (compared with $\alpha$) downstream of 40 km, indicating more homogeneous sand and gravel mixing at this longer wavelength in the foreland. A notable feature of the 10-km-binned $\beta$ trendline is the large range in IQR between 60–70-km upstream. This reach contains the coarse fan link at 65 km, upstream of which there are large sand bars (> 18 m), and downstream of which coarse gravels are frequently delivered to the channel bed during debris flow events that often damage the adjacent road and disperse gravels into the bedload. Data are sparse upstream of the gorge, largely due to a narrower channel bed with insufficient clip size for robust fitting at wavelengths > 18 m, but the spread of $\beta$ on successfully fit clips is large in this reach, where larger sand bars with limited contiguous gravel patches are locally formed in 1-km reaches.

# Effect of Polarization

For all analyses to this point we have only used the dual-polarized bands (HH for TSX/TDX and ALOS2 or VV for S1). Differences in HH and VV polarization for non-vegetated surfaces are likely minimal, particularly since the gravel elements cause intensity scattering in both the vertical and horizontal directions, but cross-polarized results show some difference. For the S1 scenes, we note similar inside-channel trends between the VV dual- and VH cross-polarized bands in Figures C8A and C9A, respectively. In the S1 VH cross-polarized outside-channel trendline, there is a near match with the inside-channel trend (Fig. C9A), which leads to nearly flat to slightly decreasing $\sigma^0$ for the S1 VH corrected trendline (Fig. C9B). This difference of the S1 VH band outside-channel trend is likely the result of vegetation (only with NDVI < 0.35), which causes modulation of the cross-polarized signal by volume-scattering effects (e.g., Vreugdenhil et al., 2018). Close-packed boulders have also been shown to exhibit a similar response (Evans et al., 1986), but large boulders in the Río Toro are sparsely distributed. In contrast to the S1 C-band data, the ALOS2 L-band HV cross-polarized trend (Fig. C11B) is comparable (only lower in magnitude) with

the ALOS2 HH dual-polarized trend (Fig. C10B). This is likely because the longer wavelength is able to penetrate the remaining unmasked sparse vegetation, leading to a similar response between both polarizations.

## Caveats and Application of the Method

In the present study, we have not considered radar penetration into dry sand (e.g., McCauley et al., 1982). We are less concerned with this issue since even very low moisture content has been shown to block radar penetration after only a few centimeters, even for long wavelengths (Williams and Greeley, 2004). In our channel bed, such low moisture contents are likely present throughout the reach (Bookhagen and Strecker, 2008; Castino et al., 2016b), so SAR penetration should be restricted to only a thin layer of the smooth sand surfaces, which are typically centimeters- to meters-thick deposits. Although SAR amplitude is sensitive to moisture as well as roughness, most studies have covered agricultural soils made up of silt and clay (e.g., Baghdadi et al., 2008), where water retention and soil moisture is much higher than granular sand, so remaining effects of moisture following our correction steps are likely minimal, and dominated by the surface roughness response.

In highly active alluvial channels in mountain environments with mixed sand and gravel bars, the patchiness wavelength often occurs over short ($\leq 50$ m) distances. Thus, our TSX/TDX data are likely near the limit of sand-gravel patchiness measurements, with lower-resolution data (e.g., 15 m) integrating signals over scales at which the channel bed is more homogeneous. Higher-resolution TSX/TDX data are possible from the spotlight sensing mode (e.g., Aubert et al., 2011) or with lower multilooking, however, there is a trade off between the signal-to-noise ratio as resolution or multilooking is decreased. Channel-bed patchiness via the power spectrum approach is likely limited to only high-resolution SAR sensors, but it is nevertheless a novel use of DFT analysis for SAR amplitude data, extendable to other surfaces of interest (e.g., agricultural fields).

Mapping full grain-size distributions is challenging from SAR data alone. Longer-wavelength P-band ($\lambda = 68$ cm) measurements have shown even greater ranges in roughness sensitivity with some potential for mapping grain-size distributions using combinations of L-, and P-band (e.g., Campbell and Shepard, 1996; Campbell, 2001), but these sensors are rare. Thus, for the time being, SAR amplitude measurements over alluvial sand- and gravel-bed rivers in high-mountain environments are limited to measuring the gravel-to-sand transition zones. This transition is of great interest in understanding mechanisms of downstream fining via processes of abrasion or selective transport (Lamb and Venditti, 2016; Dingle et al., 2017), and the proposed methods could be applied at large scales using river outlines extracted by manual digitization on freely available imagery (Fisher et al., 2013).

# Conclusions

SAR amplitude is a useful remote-sensing tool for measuring surface roughness from space. Previous research has focused primarily on applications to water surfaces, agricultural fields, or dry desert alluvial fans. Here, we have presented a novel use of SAR amplitude data applied to high-mountain alluvial channel bedload measurements. Care must be taken in mitigating the effects of vegetation and soil moisture on these measurements. The described method is not able to map full grain-size distributions, but the transition between radar-smooth sand surfaces and radar-rough gravels can be assessed, and trends in these bedload arrangements are visible. Longer wavelength SAR (e.g., L-band) shows a larger range of measurement, and higher spatial resolution SAR (e.g., TerraSAR-

X/TanDEM-X) can be used for detailed geomorphic characterization. Mapping the sand-gravel transitions in alluvial rivers can be done using even short wavelength and coarse spatial-resolution SAR sensors. The methods presented here increase the breadth of environmental measurements possible from spaceborne radar. The spatial scales of these observations can be greatly increased to entire orogenic belts, and time series, such as those regularly generated from Sentinel-1, can be used to look at temporal dynamics of sand and gravel bedload.

# Acknowledgments

The authors thank the DLR for TerraSAR-X/TanDEM-X scenes received through proposal XTI_GEOL6727 to B. Bookhagen. Project funding was sourced from DFG funded IRTG-StRATEGy (IGK2018) and NEXUS funded through the MWFK Brandenburg, Germany, both for B. Bookhagen.

# Supporting Information

See Appendix C.

Chapter 6

# Discussion, Outlook, and Conclusions

Increasingly, remote sensing is being exploited in the measurement of earth surface processes. These technologies, and their associated techniques, require careful consideration to mitigate errors, especially in the high mountains. The objective of this dissertation was to answer geomorphic questions using remote sensing data to upscale previously localized results. Gathering detailed observations and physical samples of point measurements via traditional fieldwork provides the lowest error in data collection. The draw of remote sensing is the ability to quickly gather much larger datasets at the scale of entire catchments or orogens, but with the risk of higher uncertainties and misinterpretation of results.

The strong gradients and diversity of geomorphic expression in the study area allowed this dissertation to assess a broad array of near (camera-based) and far (spaceborne DEMs and SAR) remote sensing data, and to advance methods of their application. All three published or submitted studies focus on the realm of high-mountain fluvial geomorphology, with the exception of landscape-wide measurements explored in Chapter 3. These low-slope, non-vegetated alluvial rivers are ideal environments for remote sensing of bare-earth surface processes, where aerially exposed bedload can be measured without complicating factors of rugged topography or landcover by dense forest. The studies contained in this dissertation demonstrate the potential for a new generation of earth surface processes scientists to draw on quantitative analysis of remote sensing data to achieve unprecedented scales of traditional measurements of the mechanisms, rates, and magnitudes of geomorphic expression. In the following sections, the results of Chapters 3–5 are summarized in the context of the guiding research questions, future outlooks from each technique are outlined, and concluding messages are drawn from this dissertation.

## Research Question 1: Chapter 3

### *What are the limits of global-scale spaceborne DEM differencing for measuring land-level changes in high-relief terrain?*

Monitoring land-level changes over annual to decadal time scales requires some combination of high-precision measurements and large rates and magnitudes of change. High-mountain rivers present an ideal, low-slope setting to explore less-accurate spaceborne data applied to this question. Given the typically low (< 1 m/yr) rates of height changes on the ice- and snow-free terrestrial earth, measuring these rates often requires very high-precision data. These data are usually attained from near remote sensing surveys on the ground (e.g., Wheaton et al., 2010) or

from airplanes or drones, which can be equipped with lidar (e.g., Lane et al., 2003; Mason and Mohrig, 2018), or cameras for traditional photogrammetry or SfM (e.g., Brasington et al., 2003; Cook, 2017).

Field observations of buried roads, aggradation under bridges, and preventative gravel mining in the Eastern Central Andes indicate high rates (> 1 m/yr) of change in local river segments. Furthermore, the occurrence of frequent landslides in this region (Savi et al., 2016; Olen and Bookhagen, 2018) contributes to localized land-level changes with tens- to hundreds-of-meter magnitudes. There is thus high potential for land-level change monitoring using less precise spaceborne data, typically reserved for cryospheric studies of ice and snow changes (e.g., Brun et al., 2017; Neelmeijer et al., 2017; Dussaillant et al., 2019). Although higher-resolution, more accurate DEMs can be gathered at the catchment scale from such optical satellites as Pleiades or WorldView, these data remain costly. With newly generated DEMs from TanDEM-X, there is now a large time step (~15 years) between global-scale spaceborne DEMs. In Chapter 3, this dissertation seeks to exploit this time step and the high rates of expected change to monitor land-level changes at larger scales than previously possible using more expensive or small-areal remote sensing data and techniques.

The study achieved this analysis through the development of a semi-automated pipeline for bias correction and the identification of pixels with vertical differences beyond the statistical range of uncertainty in measurement. In general, this study found that despite recent re-processing of the original SRTM data (Crippen et al., 2016), land-level change monitoring is hampered by the lower quality of older DEMs. This is especially true on steeper slopes, where even newer TanDEM-X data suffered from high uncertainties, which limit the application of spaceborne DEM land-level change monitoring to relatively flat, low-slope surfaces. Measurement is further limited by the spatial resolution of spaceborne DEMs (30 m the case of those used in Chapter 3), whereas point cloud data from lidar or SfM can achieve much higher sub-meter resolution measurements. Finally, it is suggested that spaceborne DEMs generated via radar interferometry offer more accurate representation of the earth's surface compared with global optical DEMs from ASTER or ALOS (Purinton and Bookhagen, 2017).

Only one landslide could be definitively identified, as these events usually occur on steep slopes where DEM uncertainty precludes identification of error versus true geomorphic signal. Notably, outside of the fluvial environment, this study also found some auxiliary TanDEM-X data layers useful for identifying regions of change (namely, dune fields) over the ~5-year acquisition period of individual TerrSAR-X/TanDEM-X DEMs (Rizzoli et al., 2017). The rapidly aggrading and incising channel beds of the Río Toro and Río Grande allowed many more geophysical measurements of height change, as slopes are near-planar and DEM accuracy is much improved. Although natural signals of aggradation and incision were concentrated in wetter, oversteepened sections downstream of the knickpoint and near the orographic barrier, the greatest magnitude changes were shown to be anthropogenic, due to gravel-mining piles and the construction of large weirs.

## Future Outlook from Chapter 3

Overall, Chapter 3 primarily serves to highlight the biases of spaceborne DEMs as they relate to land-level change monitoring in high-relief terrain and provides a detailed account of the biases and uncertainties of one of the most widely used DEMs (SRTM-C) and the newest global DEM (TanDEM-X). Thus, the work is a road map for geomorphologists seeking to apply future spaceborne elevation measurements to terrestrial land-level changes, with

the SRTM-C and TanDEM-X remaining as year 2000 and year ~2015 snapshots, respectively. Especially enticing for such analysis are high-accuracy DEMs from spaceborne optical sensors, which can be rapidly generated using the *Ames Stereo Pipeline* (Shean et al., 2016). The sparsity of geophysical measurements found in this study should not be a deterrent for ongoing research, but rather highlights the necessity of bias correction steps and cautious interpretation of difference maps acquired from global-scale DEMs in high-mountain environments.

# Research Question 2: Chapter 4

*What are the limits of camera-based grain-size detection at the scale of entire catchments in steep and dynamic gravel-bed rivers in the high mountains?*

Traditional measurement of fluvial grain-size distributions by physical counting is difficult at the scale of entire catchments, and data tends to be sparse (e.g., Dingle et al., 2017). Difficulty is further increased in high-mountain river systems with a large range in grain size and diverse arrangements of pebble and sand patches, making representative sample gathering challenging (Wolcott and Church, 1991). Many digital grain-sizing techniques have been developed on lower-energy fluvial environments with oblate grains and smaller ranges in grain size (Butler et al., 2001; Sime and Ferguson, 2003; Graham et al., 2005a,b; Detert and Weitbrecht, 2012, 2013), but these are difficult to apply at large scales and in high-energy mountain rivers. This situation motivated the study in Chapter 4.

In a first step, a new technique for photo surveys of channel beds was developed using a camera-on-mast setup. It was found that the high resolution necessary (~1 mm/pixel) for gravel identification and sizing could not be achieved using currently available UAV technology. The use of a commercial-grade 24-megapixel camera attached to a mast with a height of 3–5 m was sufficient for capturing ~1,000–5,000 $m^2$ photo surveys at sub-cm resolution that could be processed via SfM techniques. These surveys provided both dense point clouds and seamless orthomosaics, on which new grain-sizing algorithms could be developed and applied.

Experimentation directly on the raw point-cloud data revealed the necessity of low-angle oblique imagery to properly depict the sides of grains required for 3D segmentation (see Appendix B). However, the acquisition of such photos at the scale of channel cross-sections is not feasible, and near-nadir imagery collection via the mast setup is preferred for covering large areas representative of spatially heterogeneous high-mountain rivers. Thus, the algorithms were limited to 2D image segmentation in the spectral and spatial domains from ortho-imagery. These open-source tools were published under the title *PebbleCounts* (Purinton and Bookhagen, 2019a). These algorithms are able to achieve accurate grain-size distribution measurements over large survey areas in complex fluvial environments, on the subset of grains with intermediate axis lengths $\geq$ 20 pixels ($\geq$ 2 cm for 1-mm/pixel imagery). This study also found that increasing imagery resolution, along with the acquisition of sharp, low-blur images using a high-quality lens further improved grain-size results, with the caveat of longer processing times for image segmentation.

## Future Outlook from Chapter 4

Although this study was restricted to cross-sections with a maximum size of ~5,000 $m^2$, the camera-on-mast survey technique can be easily applied at areas up to one hectare (10,000 $m^2$) in a day of work, depending on the number

of people and cameras available. Improvements in UAV technology could allow for higher-resolution surveys with lower effort in the future, but the underlying principles of SfM data generation and the necessity of high-overlap and slightly oblique (off-nadir) imagery collection will remain the same. In Chapter 4, the measurement of grain-size distributions was restricted to a number of small (~1 $m^2$) test plots from the full survey sites, on which manual identification of grain-size was carried out for control-data purposes. However, suggestions are made for processing larger sites using a combination of the automated and semi-automated algorithms presented. Thus, these tools can be applied to the full survey sites to measure grain-size distributions on tens of thousands of gravels per site, creating robust datasets for modeling downstream fining processes at unprecedented scales and data volumes in high-mountain environments (e.g., Rice and Church, 1998; Gomez et al., 2001; Chatanantavet et al., 2010; Lamb and Venditti, 2016; Dingle et al., 2017).

# Research Question 3: Chapter 5

*What is the potential of spaceborne radar for measuring bedload characteristics at the scale of entire catchments in steep and dynamic gravel-bed rivers in the high mountains?*

Similar to Chapter 4, the final study contained in this dissertation is also concerned with the measurement of grain size in high-mountain fluvial environments. Even using the advanced image processing tools contained in *PebbleCounts*, achieving continuous measurements of bedload characteristics over tens to hundreds of kilometers of channel length is not feasible. Methods exist for such analysis using optical imagery (e.g., Carbonneau et al., 2004; Mushkin and Gillespie, 2005), but these are based on correlative measures of sub-pixel shadowing, and are restricted to cloud-free imagery acquisition. On the other hand, radar data are enticing for penetrating cloud cover and directly measuring the height variations of surfaces comprised of mixed smooth sand bars and rough gravel elements.

In Chapter 5, this dissertation sought to expand radar-roughness measurements to the fluvial realm using SAR backscatter intensity. Prior work on similar geologic surfaces is limited (e.g., Deroin et al., 1997; Sadeh et al., 2018), and radar is more often used to measure the soil moisture and micro-roughness variations in agricultural fields (e.g., Baghdadi et al., 2008). Care was taken to avoid measurement contamination by vegetation and moisture contained in loose-sand bars, and three radar wavelengths (X-, C-, and L-band) were used to assess differences in response. A quantitative analysis of local geomorphic complexity was achieved by Fourier frequency analysis of SAR backscatter and by fitting logarithmic decay functions to the resulting power spectrum. It was found that longer wavelength SAR is more sensitive to grain-size variation, as shorter wavelengths saturate at even low roughness values created by gravels. Higher-resolution data is necessary for the frequency analysis, as the scales over which high-mountain channel beds change are typically at or near the resolution of currently available spaceborne SAR data. SAR amplitude-based roughness measurements were restricted to sand-gravel transitions along the studied channel, and not capable of capturing the grain-size distribution beyond this dichotomy.

## Future Outlook from Chapter 5

The extension of radar amplitude-based roughness measurements to the fluvial realm is an important development in earth surface processes research. The methods developed in this study, particularly in the frequency analysis, are

applicable to other geologic (e.g., alluvial fans) or agricultural (e.g., plowed fields) surfaces. Overall, the research contained in Chapter 5 expands the potential for orogen-scale monitoring of sand-gravel transitions (Lamb and Venditti, 2016; Dingle et al., 2017) from spaceborne remote sensing. Further work on this topic could include the generation of time series of channel bedload behavior using freely available, global-coverage Sentinel-1 radar data.

# Concluding Remarks

The objectives of this dissertation arose from traditional field observations of earth surface processes in the Eastern Central Andes. Remote sensing was the tool chosen to quantify these processes over the large spatial scales of the studied catchments. The data and methods from these studies are applicable to other rugged and steep environments, and represent important progress at the disciplinary intersection of contemporary geomorphology and remote sensing in highly dynamic mountain catchments.

In a first step, this dissertation explored *the limits of global-scale spaceborne DEM differencing for measuring land-level changes*. To this end, a differencing approach was applied to two spaceborne radar DEMs separated by ~15 years. Following this, *the limits of camera-based grain-size detection in dynamic high-mountain rivers* was explored through channel-bed photo surveys using the principles of SfM, and the development of open-source image-segmentation algorithms for grain identification and measurement. In a final study, *the potential of spaceborne radar for measuring river-bedload characteristics at the scale of entire catchments* was evaluated via SAR-amplitude roughness measurements, previously restricted to dry desert alluvial fans or agricultural fields. The following general concluding remarks are drawn from these studies:

1. Time series of DEMs are useful for measuring land-level changes, but only when the magnitudes and rates of change exceed expected uncertainties. For global-scale spaceborne DEMs, this constraint limits measurements to lower-slope channel-bed aggradation and incision in regions with high mass-transfer rates, perhaps influenced by anthropogenic tampering. Steeper, landslide-prone hillslopes are more difficult to measure with relatively low-accuracy globally available spaceborne DEMs. (Chapter 3)

2. Although open-access spaceborne DEMs have lower accuracy compared to purchased and/or manually generated data (e.g., from high-resolution optical satellites, and aerial or terrestrial lidar or SfM), these datasets will remain as snapshots of the earth's surface at the time of collection, and are therefore useful reference surfaces in future land-level change studies. (Chapter 3)

3. The recent proliferation of UAV-based surveying does not preclude handheld or mast-mounted camera-based techniques. Commercial cameras can achieve higher resolution results at low cost for monitoring fluvial systems. (Chapter 4)

4. Point-cloud segmentation of channel-bed surveys require not only the tops of grains from near-nadir imagery, but also their sides, which requires many more photos from oblique viewing angles. Thus, channel-bed surveys using only near-nadir imagery are restricted to 2D image-segmentation algorithms. (Chapter 4)

5. SAR amplitude measurements can be applied to measure channel bedload characteristics in alluvial rivers with aerially exposed sand and gravel bars. However, the saturation of backscatter intensity on even low-roughness surfaces restricts these measurements to sand-gravel transitions. (Chapter 5)

6. Uncertainties related to sensor biases and complexity of the natural environment must be rigorously assessed prior to interpretation of remote sensing results for earth surface processes research in the high mountains, where steep topography and strong gradients in geomorphic expression can complicate measurements. (Chapters 3, 4, and 5)

7. In addition to the accuracy of measurements, their spatial resolution is a limiting factor when applied to earth surface processes, with higher resolutions allowing better characterization of local geomorphic processes, magnitudes, and rates. (Chapters 3, 4, and 5)

8. As the spatial resolution, temporal resolution, and accuracy of near and far remote sensing data increases — and their analysis techniques become more refined — the use of remote sensing in measuring earth surface processes will continue to bridge gaps between plot- and catchment-scale measurements in high-mountain environments, towards the goal of synoptic, orogen-scale monitoring of landscape form and rates of change.

# References


Agisoft: AgiSoft MetaShape Professional, URL http://www.agisoft.com/downloads/installer/, 2019.

Ahnert, F.: Brief description of a comprehensive three-dimensional process-response model for landform development, Zeitschrift fur Geomorphologie, NF, Supplement, 25, 29–49, 1976.

Allmendinger, R. W., Jordan, T. E., Kay, S. M., and Isacks, B. L.: The Evolution of the Altiplano-Puna Plateau of the Central Andes, Annual Review of Earth and Planetary Sciences, 25, 139–174, doi: 10.1146/annurev.earth.25.1.139, 1997.

Alonzo, M., Bookhagen, B., McFadden, J. P., Sun, A., and Roberts, D. A.: Mapping urban forest leaf area index with airborne lidar using penetration metrics and allometry, Remote Sensing of Environment, 162, 141–153, doi: 10.1016/j.rse.2015.02.025, 2015.

Arrell, K., Wise, S., Wood, J., and Donoghue, D.: Spectral filtering as a method of visualising and removing striped artefacts in digital elevation data, Earth Surface Processes and Landforms, 33, 943–961, 2008.

Attal, M. and Lavé, J.: Changes of bedload characteristics along the Marsyandi River (central Nepal): Implications for understanding hillslope sediment supply, sediment load evolution along fluvial networks, and denudation in active orogenic belts, Geological Society of America Special Papers, 398, 143–171, doi: 10.1130/2006.2398(09), 2006.

Attal, M., Mudd, S., Hurst, M., Weinman, B., Yoo, K., and Naylor, M.: Impact of change in erosion rate and landscape steepness on hillslope and fluvial sediments grain size in the Feather River basin (Sierra Nevada, California), Earth Surface Dynamics, 3, 201–222, doi: 10.5194/esurf-3-201-2015, 2015.

Aubert, M., Baghdadi, N., Zribi, M., Douaoui, A., Loumagne, C., Baup, F., Hajj, M. E., and Garrigues, S.: Analysis of TerraSAR-X data sensitivity to bare soil moisture, roughness, composition and soil crust, Remote Sensing of Environment, 115, 1801 – 1810, doi: https://doi.org/10.1016/j.rse.2011.02.021, URL http://www.sciencedirect.com/science/article/pii/S003442571100071X, 2011.

Baade, J. and Schmullius, C.: TanDEM-X IDEM precision and accuracy assessment based on a large assembly of differential GNSS measurements in Kruger National Park, South Africa, ISPRS Journal of Photogrammetry and Remote Sensing, 119, 496–508, 2016.

Baade, J. and Schmullius, C. C.: Interferometric Microrelief Sensing With TerraSAR-X—First Results, IEEE Transactions on Geoscience and Remote Sensing, 48, 965–970, doi: 10.1109/TGRS.2009.2036720, 2010.

Baghdadi, N. and Zribi, M.: Evaluation of radar backscatter models IEM, OH and Dubois using experimental observations, International Journal of Remote Sensing, 27, 3831–3852, doi: 10.1080/01431160600658123, URL https://doi.org/10.1080/01431160600658123, 2006.

Baghdadi, N., Zribi, M., Loumagne, C., Ansart, P., and Anguela, T. P.: Analysis of TerraSAR-X data and their sensitivity to soil surface parameters over bare agricultural fields, Remote Sensing of Environment, 112, 4370 – 4379, doi: https://doi.org/10.1016/j.rse.2008.08.004, URL http://www.sciencedirect.com/science/article/pii/S0034425708002496, 2008.

Baghdadi, N., El Hajj, M., Choker, M., Zribi, M., Bazzi, H., Vaudour, E., Gilliot, J.-M., and Ebengo, D. M.: Potential of Sentinel-1 Images for Estimating the Soil Roughness over Bare Agricultural Soils, Water, 10, doi: 10.3390/w10020131, URL https://www.mdpi.com/2073-4441/10/2/131, 2018.

Bagnardi, M., González, P. J., and Hooper, A.: High-resolution digital elevation model from tri-stereo Pleiades-1 satellite imagery for lava flow volume estimates at Fogo Volcano, Geophysical Research Letters, 43, 6267–6275, doi: 10.1002/2016gl069457, 2016.

Barnhart, W. D., Gold, R. D., Shea, H. N., Peterson, K. E., Briggs, R. W., and Harbor, D. J.: Vertical Coseismic Offsets Derived From High-Resolution Stereogrammetric DSM Differencing: The 2013 Baluchistan, Pakistan Earthquake, Journal of Geophysical Research: Solid Earth, 124, 6039–6055, doi: 10.1029/2018JB017107, 2019.

Berthier, E., Arnaud, Y., Vincent, C., and Rémy, F.: Biases of SRTM in high-mountain areas: Implications for the monitoring of glacier volume changes, Geophysical Research Letters, 33, doi: 10.1029/2006gl025862, 2006.

Berthier, E., Arnaud, Y., Kumar, R., Ahmad, S., Wagnon, P., and Chevallier, P.: Remote sensing estimates of glacier mass balances in the Himachal Pradesh (Western Himalaya, India), Remote Sensing of Environment, 108, 327–338, doi: 10.1016/j.rse.2006.11.017, 2007.

Bertin, S. and Friedrich, H.: Field application of close-range digital photogrammetry (CRDP) for grain-scale fluvial

morphology studies, Earth Surface Processes and Landforms, 41, 1358–1369, doi: 10.1002/esp.3906, 2016.

Bertin, S., Groom, J., and Friedrich, H.: Isolating roughness scales of gravel-bed patches, Water Resources Research, 53, 6841–6856, doi: 10.1002/2016WR020205, 2017.

Bessette-Kirton, E. K., Coe, J. A., and Zhou, W.: Using Stereo Satellite Imagery to Account for Ablation, Entrainment, and Compaction in Volume Calculations for Rock Avalanches on Glaciers: Application to the 2016 Lamplugh Rock Avalanche in Glacier Bay National Park, Alaska, Journal of Geophysical Research: Earth Surface, 123, 622–641, doi: 10.1002/2017JF004512, 2018.

Bierman, P. and Steig, E. J.: Estimating rates of denudation using cosmogenic isotope abundances in sediment, Earth Surface Processes and Landforms, 21, 125–139, doi: 10.1002/(SICI)1096-9837(199602)21:2<125::AID-ESP511>3.0.CO;2-8, 1996.

Bookhagen, B. and Burbank, D. W.: Toward a complete Himalayan hydrological budget: Spatiotemporal distribution of snowmelt and rainfall and their impact on river discharge, Journal of Geophysical Research: Earth Surface, 115, doi: 10.1029/2009JF001426, 2010.

Bookhagen, B. and Strecker, M. R.: Orographic barriers, high-resolution TRMM rainfall, and relief variations along the eastern Andes, Geophysical Research Letters, 35, doi: 10.1029/2007gl032011, 2008.

Bookhagen, B. and Strecker, M. R.: Spatiotemporal trends in erosion rates across a pronounced rainfall gradient: Examples from the southern Central Andes, Earth and Planetary Science Letters, 327-328, 97–110, doi: 10.1016/j.epsl.2012.02.005, 2012.

Booth, A. M., Roering, J. J., and Perron, J. T.: Automated landslide mapping using spectral analysis and high-resolution topographic data: Puget Sound lowlands, Washington, and Portland Hills, Oregon, Geomorphology, 109, 132–147, doi: 10.1016/j.geomorph.2009.02.027, 2009.

Bourgine, B. and Baghdadi, N.: Assessment of C-band SRTM DEM in a dense equatorial forest zone, Comptes Rendus Geoscience, 337, 1225–1234, 2005.

Bousbih, S., Zribi, M., Lili-Chabaane, Z., Baghdadi, N., El Hajj, M., Gao, Q., and Mougenot, B.: Potential of Sentinel-1 Radar Data for the Assessment of Soil and Cereal Cover Parameters, Sensors, 17, doi: 10.3390/s17112617, URL https://www.mdpi.com/1424-8220/17/11/2617, 2017.

Bradski, G.: The OpenCV Library, 2000.

Brasington, J., Langham, J., and Rumsby, B.: Methodological sensitivity of morphometric estimates of coarse fluvial sediment transport, Geomorphology, 53, 299–316, 2003.

Brasington, J., Vericat, D., and Rychkov, I.: Modeling river bed morphology, roughness, and surface sedimentology using high resolution terrestrial laser scanning, Water Resources Research, 48, W11 519, doi: 10.1029/2012WR012223, 2012.

Brun, F., Berthier, E., Wagnon, P., Kääb, A., and Treichler, D.: A spatially resolved estimate of High Mountain Asia glacier mass balances from 2000 to 2016, Nature Geosci, advance online publication, doi: 10.1038/ngeo2999, 2017.

Bryant, R., Moran, M. S., Thoma, D. P., Holifield Collins, C. D., Skirvin, S., Rahman, M., Slocum, K., Starks, P., Bosch, D., and Gonzalez Dugo, M. P.: Measuring Surface Roughness Height to Parameterize Radar Backscatter Models for Retrieval of Surface Soil Moisture, IEEE Geoscience and Remote Sensing Letters, 4, 137–141, doi: 10.1109/LGRS.2006.887146, 2007.

Buades, A., Coll, B., and Morel, J.-M.: Non-Local Means Denoising, Image Processing On Line, 1, 208–212, doi: 10.5201/ipol.2011.bcm_nlm, 2011.

Bunte, K. and Abt, S. T.: Sampling surface and subsurface particle-size distributions in wadable gravel- and cobble-bed streams for analyses in sediment transport, hydraulics and streambed monitoring, Tech. rep., US Forest Service, Rocky Mountain Research Station, Fort Collins, CO, doi: 10.2737/RMRS-GTR-74, 2001.

Burbank, D. W., Leland, J., Fielding, E., Anderson, R. S., Brozovic, N., Reid, M. R., and Duncan, C.: Bedrock incision, rock uplift and threshold hillslopes in the northwestern Himalayas, Nature, 379, 505, doi: 10.1038/379505a0, 1996.

Bürgmann, R., Hilley, G., Ferretti, A., and Novali, F.: Resolving vertical tectonics in the San Francisco Bay Area from permanent scatterer InSAR and GPS analysis, Geology, 34, 221–224, doi: 10.1130/G22064.1, 2006.

Buscombe, D.: Transferable wavelet method for grain-size distribution from images of sediment surfaces and thin sections, and other natural granular patterns, Sedimentology, 60, 1709–1732, doi: 10.1111/sed.12049, 2013.

Buscombe, D., Rubin, D. M., and Warrick, J. A.: A universal approximation of grain size from images of noncohesive sediment, Journal of Geophysical Research: Earth Surface, 115, F02 015, doi: 10.1029/2009JF001477, 2010.

Butler, J. B., Lane, S. N., and Chandler, J. H.: Automated extraction of grain-size data from gravel surfaces using digital image processing, Journal of Hydraulic Research, 39, 519–529, doi: 10.1080/00221686.2001.9628276, 2001.

Campbell, B. A.: Radar Backscatter from Mars: Properties of Rock-Strewn Surfaces, Icarus, 150, 38 – 47, doi: https://doi.org/10.1006/icar.2000.6566, URL http://www.sciencedirect.com/science/article/pii/S0019103500965667, 2001.

Campbell, B. A. and Shepard, M. K.: Lava flow surface roughness and depolarized radar scattering, Journal of Geophysical Research: Planets, 101, 18 941–18 951, doi: 10.1029/95JE01804, URL https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/95JE01804, 1996.

Carabajal, C. C. and Harding, D. J.: SRTM C-band and ICESat laser altimetry elevation comparisons as a function of tree cover and relief, Photogrammetric Engineering & Remote Sensing, 72, 287–298, 2006.

Carbonneau, P., Bizzi, S., and Marchetti, G.: Robotic photosieving from low-cost multirotor sUAS: a proof-of-concept, Earth Surface Processes and Landforms, 43, 1160–1166, doi: 10.1002/esp.4298, 2018.

Carbonneau, P. E.: The threshold effect of image resolution on image-based automated grain size mapping in fluvial environments, Earth Surface Processes and Landforms, 30, 1687–1693, doi: 10.1002/esp.1288, 2005.

Carbonneau, P. E. and Dietrich, J. T.: Cost-effective non-metric photogrammetry from consumer-grade sUAS: implications for direct georeferencing of structure from motion photogrammetry, Earth Surface Processes and Landforms, 42, 473–486, doi: doi:10.1002/esp.4012, 2017.

Carbonneau, P. E., Lane, S. N., and Bergeron, N. E.: Cost-effective non-metric close-range digital photogrammetry and its application to a study of coarse gravel river beds, International Journal of Remote Sensing, 24, 2837–2854, doi: 10.1080/01431160110108364, 2003.

Carbonneau, P. E., Lane, S. N., and Bergeron, N. E.: Catchment-scale mapping of surface grain size in gravel bed rivers using airborne digital imagery, Water Resources Research, 40, W07 202, doi: 10.1029/2003WR002759, 2004.

Castino, F., Bookhagen, B., and Strecker, M.: River-discharge dynamics in the Southern Central Andes and the 1976–77 global climate shift, Geophysical Research Letters, 43, doi: 10.1002/2016GL070868, 2016a.

Castino, F., Bookhagen, B., and Strecker, M.: Rainfall variability and trends of the past six decades (1950–2014) in the subtropical NW Argentine Andes, Climate Dynamics, 3, 1049–1067, 2016b.

Castino, F., Bookhagen, B., and Strecker, M. R.: Oscillations and trends of river discharge in the southern Central Andes and linkages with climate variability, Journal of Hydrology, 555, 108–124, doi: 10.1016/j.jhydrol.2017.10.001, 2017.

Chatanantavet, P., Lajeunesse, E., Parker, G., Malverti, L., and Meunier, P.: Physically based model of downstream fining in bedrock streams with lateral input, Water Resources Research, 46, W02 518, doi: 10.1029/2008WR007208, 2010.

Chen, Q., Baldocchi, D., Gong, P., and Kelly, M.: Isolating Individual Trees in a Savanna Woodland Using Small Footprint Lidar Data, Photogrammetric Engineering & Remote Sensing, 72, 923–932, doi: 10.14358/PERS.72.8.923, 2006.

Church, M.: Geomorphic thresholds in riverine landscapes, Freshwater biology, 47, 541–557, 2002.

Church, M., Hassan, M. A., and Wolcott, J. F.: Stabilizing self-organized structures in gravel-bed stream channels: Field and experimental observations, Water Resources Research, 34, 3169–3179, doi: 10.1029/98WR00484, 1998.

Clauset, A., Shalizi, C. R., and Newman, M. E.: Power-law distributions in empirical data, SIAM review, 51, 661–703, doi: 10.1137/070710111, 2009.

CloudCompare: CloudCompare Software, URL http://www.cloudcompare.org/, 2018.

Clubb, F. J., Mudd, S. M., Attal, M., Milodowski, D. T., and Grieve, S. W.: The relationship between drainage density, erosion rate, and hilltop curvature: Implications for sediment transport processes, Journal of Geophysical Research: Earth Surface, 121, 1724–1745, 2016.

Comiti, F., Da Canal, M., Surian, N., Mao, L., Picco, L., and Lenzi, M.: Channel adjustments and vegetation cover dynamics in a large gravel bed river over the last 200 years, Geomorphology, 125, 147–159, 2011.

Cook, K. L.: An evaluation of the effectiveness of low-cost UAVs and structure from motion for geomorphic change detection, Geomorphology, 278, 195–208, 2017.

Crippen, R., Buckley, S., Agram, P., Belz, E., Gurrola, E., Hensley, S., Kobrick, M., Lavalle, M., Martin, J., Neumann, M., et al.: NASADEM global elevation model: methods and progress, ISPRS-International Archives of the Photogrammetry Remote Sensing and Spatial Information Sciences, XLI-B4, 125–128, doi: 10.5194/isprs-archives-XLI-B4-125-2016, 2016.

Cullen, N. D., Verma, A. K., and Bourke, M. C.: A comparison of structure from motion photogrammetry and the traversing micro-erosion meter for measuring erosion on shore platforms, Earth Surface Dynamics, 6, 1023–1039, doi: 10.5194/esurf-6-1023-2018, URL https://www.earth-surf-dynam.net/6/1023/2018/, 2018.

de Haas, T., Ventra, D., Carbonneau, P. E., and Kleinhans, M. G.: Debris-flow dominance of alluvial fans masked by runoff reworking and weathering, Geomorphology, 217, 165 – 181, doi: 10.1016/j.geomorph.2014.04.028, 2014.

Deroin, J. ., Company, A., and Simonin, A.: An empirical model for interpreting the relationship between backscattering and arid land surface roughness as seen with the SAR, IEEE Transactions on Geoscience and Remote Sensing, 35, 86–92, doi: 10.1109/36.551937, 1997.

Detert, M. and Weitbrecht, V.: Automatic object detection to analyze the geometry of gravel grains–a free stand-alone tool, in: River flow 2012 : Proceedings of the international conference on fluvial hydraulics, San José, Costa Rica, September 5-7, 2012, pp. 595–600, Taylor & Francis Group, London, 2012.

Detert, M. and Weitbrecht, V.: User guide to gravelometric image analysis by BASEGRAIN, pp. 1789–1796, Advances in River Sediment Research, 2013 Taylor & Francis Group, London, 2013.

DiBiase, R. A., Whipple, K. X., Heimsath, A. M., and Ouimet, W. B.: Landscape form and millennial erosion rates in the San Gabriel Mountains, CA, Earth and Planetary Science Letters, 289, 134–144, doi: 10.1016/j.epsl.2009.10.036, 2010.

Dierking, W.: Quantitative roughness characterization of geological surfaces and implications for radar signature analysis, IEEE Transactions on Geoscience and Remote Sensing, 37, 2397–2412, doi: 10.1109/36.789638, 1999.

Dietrich, W. E. and Perron, J. T.: The search for a topographic signature of life, Nature, 439, 411, doi: 10.1038/nature04452, 2006.

Dietrich, W. E., Bellugi, D. G., Sklar, L. S., Stock, J. D., Heimsath, A. M., and Roering, J. J.: Geomorphic Transport Laws for Predicting Landscape Form and Dynamics, in: Prediction in Geomorphology, vol. 135, pp. 103–132, Wiley-Blackwell, doi: 10.1029/135gm09, 2003.

Dingle, E. H., Attal, M., and Sinclair, H. D.: Abrasion-set limits on Himalayan gravel flux, Nature, 544, 471–474, 2017.

Dodds, P. S. and Rothman, D. H.: Scaling, Universality, and Geomorphology, Annual Review of Earth and Planetary Sciences, 28, 571–610, doi: 10.1146/annurev.earth.28.1.571, URL https://doi.org/10.1146/annurev.earth.28.1.571, 2000.

Dugdale, S. J., Carbonneau, P. E., and Campbell, D.: Aerial photosieving of exposed gravel bars for the rapid calibration of airborne grain size maps, Earth Surface Processes and Landforms, 35, 627–639, doi: 10.1002/esp.1936, 2010.

Dunne, K. B. and Jerolmack, D. J.: Evidence of, and a proposed explanation for, bimodal transport states in alluvial rivers, Earth Surface Dynamics, 6, 583–594, doi: 10.5194/esurf-6-583-2018, 2018.

Dussaillant, I., Berthier, E., Brun, F., Masiokas, M., Hugonnet, R., Favier, V., Rabatel, A., Pitte, P., and Ruiz, L.: Two decades of glacier mass loss along the Andes, Nature Geoscience, pp. 1–7, doi: 10.1038/s41561-019-0432-5, 2019.

Eltner, A., Kaiser, A., Castillo, C., Rock, G., Neugirg, F., and Abellán, A.: Image-based surface reconstruction in geomorphometry – merits, limits and developments, Earth Surface Dynamics, 4, 359–389, doi: 10.5194/esurf-4-359-2016, 2016.

Evans, D. L., Farr, T. G., Ford, J. P., Thompson, T. W., and Werner, C. L.: Multipolarization Radar Images for Geologic Mapping and Vegetation Discrimination, IEEE Transactions on Geoscience and Remote Sensing, GE-24, 246–257, doi: 10.1109/TGRS.1986.289644, 1986.

Evans, D. L., Farr, T. G., and van Zyl, J. J.: Estimates of surface roughness derived from synthetic aperture radar (SAR) data, IEEE Transactions on Geoscience and Remote Sensing, 30, 382–389, doi: 10.1109/36.134087, 1992.

Farr, T.: Guide to Magellan Image Interpretation, chap. Chapter 5: Radar Interactions with Geologic Surfaces, Jet Propulsion Laboratory, California Institute of Technology: La Cañada Flintridge, CA, USA, 1993.

Farr, T. G. and Chadwick, O. A.: Geomorphic processes and remote sensing signatures of alluvial fans in the Kun Lun Mountains, China, Journal of Geophysical Research: Planets, 101, 23 091–23 100, doi: 10.1029/96JE01603, URL https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/96JE01603, 1996.

Farr, T. G., Rosen, P. A., Caro, E., Crippen, R., Duren, R., Hensley, S., Kobrick, M., Paller, M., Rodriguez, E., Roth, L., Seal, D., Shaffer, S., Shimada, J., Umland, J., Werner, M., Oskin, M., Burbank, D., and Alsdorf, D.: The Shuttle Radar Topography Mission, Reviews of Geophysics, 45, doi: 10.1029/2005RG000183, 2007.

Ferguson, R., Hoey, T., Wathen, S., and Werritty, A.: Field evidence for rapid downstream fining of river gravels through selective transport, Geology, 24, 179–182, doi: 10.1130/0091-7613(1996)024<0179:FEFRDF>2.3.CO;2, 1996.

Fernandez Luque, R. and Van Beek, R.: Erosion and transport of bed-load sediment, Journal of hydraulic research, 14, 127–144, doi: 10.1080/00221687609499677, 1976.

Fisher, G. B., Bookhagen, B., and Amos, C. B.: Channel planform geometry and slopes from freely available high-spatial resolution imagery and DEM fusion: Implications for channel width scalings, erosion proxies, and fluvial signatures in tectonically active landscapes, Geomorphology, 194, 46–56, doi: 10.1016/j.geomorph.2013.04.011, 2013.

Fonstad, M. A., Dietrich, J. T., Courville, B. C., Jensen, J. L., and Carbonneau, P. E.: Topographic structure from motion: a new development in photogrammetric measurement, Earth Surface Processes and Landforms, 38, 421–430, doi: 10.1002/esp.3366, 2013.

Forte, A. M., Whipple, K. X., Bookhagen, B., and Rossi, M. W.: Decoupling of modern shortening rates, climate, and topography in the Caucasus, Earth and Planetary Science Letters, 449, 282–294, doi: 10.1016/j.epsl.2016.06.013, 2016.

Fripp, J. B. and Diplas, P.: Surface Sampling in Gravel Streams, Journal of Hydraulic Engineering, 119, 473–490, doi: 10.1061/(ASCE)0733-9429(1993)119:4(473), 1993.

Gallant, J. and Read, A.: Enhancing the SRTM data for Australia, Proc. Geomorphometry, 31, 149–154, 2009.

Gardelle, J., Berthier, E., and Arnaud, Y.: Impact of resolution and radar penetration on glacier elevation changes computed from DEM differencing, Journal of Glaciology, 58, 419–422, doi: 10.3189/2012JoG11J175, 2012.

GDAL/OGR contributors: GDAL/OGR Geospatial Data Abstraction software Library, Open Source Geospatial Foundation, URL https://gdal.org, 2019.

Gilbert, G. K.: Report on the Geology of the Henry Mountains, Tech. rep., 1877.

Girod, L., Nuth, C., Kääb, A., McNabb, R., and Galland, O.: MMASTER: Improved ASTER DEMs for Elevation Change Monitoring, Remote Sensing, 9, doi: 10.3390/rs9070704, URL http://www.mdpi.com/2072-4292/9/7/704, 2017.

Gomez, B., Rosser, B. J., Peacock, D. H., Hicks, D. M., and Palmer, J. A.: Downstream fining in a rapidly aggrading gravel bed river, Water Resources Research, 37, 1813–1823, doi: 10.1029/2001WR900007, 2001.

Gorokhovich, Y. and Voustianiouk, A.: Accuracy assessment of the processed SRTM-based elevation data by CGIAR using field data from USA and Thailand and its relation to the terrain characteristics, Remote Sensing of Environment, 104, 409–415, doi: 10.1016/j.rse.2006.05.012, 2006.

Gorrab, A., Zribi, M., Baghdadi, N., Mougenot, B., and Chabaane, Z. L.: Potential of X-Band TerraSAR-X and COSMO-SkyMed SAR Data for the Assessment of Physical Soil Parameters, Remote Sensing, 7, 747–766, doi: 10.3390/rs70100747, URL https://www.mdpi.com/2072-4292/7/1/747, 2015.

Graham, D. J., Reid, I., and Rice, S. P.: Automated Sizing of Coarse-Grained Sediments: Image-Processing Procedures, Mathematical Geology, 37, 1–28, doi: 10.1007/s11004-005-8745-x, 2005a.

Graham, D. J., Rice, S. P., and Reid, I.: A transferable method for the automated grain sizing of river gravels, Water Resources Research, 41, W07 020, doi: 10.1029/2004WR003868, 2005b.

Graham, D. J., Rollet, A.-J., Piégay, H., and Rice, S. P.: Maximizing the accuracy of image-based surface sediment sampling techniques, Water Resources Research, 46, W02 508, doi: 10.1029/2008WR006940, 2010.

Grant, G. E.: The Geomorphic Response of Gravel-Bed Rivers to Dams: Perspectives and Prospects, chap. 15, pp. 165–181, Wiley-Blackwell, doi: 10.1002/9781119952497.ch15, 2012.

Grieve, S. W. D., Mudd, S. M., Milodowski, D. T., Clubb, F. J., and Furbish, D. J.: How does grid-resolution modulate the topographic expression of geomorphic processes?, Earth Surface Dynamics, 4, 627–653, doi: 10.5194/esurf-4-627-2016, 2016.

Hack, J. T.: Studies of longitudinal stream profiles in Virginia and Maryland, Tech. rep., 1957.

Haralick, R. M., Shanmugam, K., and Dinstein, I.: Textural Features for Image Classification, IEEE Transactions on Systems, Man, and Cybernetics, SMC-3, 610–621, doi: 10.1109/TSMC.1973.4309314, 1973.

Heimsath, A. M., Dietrich, W. E., Nishiizumi, K., and Finkel, R. C.: The soil production function and landscape equilibrium, Nature, 388, 358, doi: 10.1038/41056, 1997.

Hetz, G., Mushkin, A., Blumberg, D. G., Baer, G., and Ginat, H.: Estimating the age of desert alluvial surfaces with spaceborne radar data, Remote Sensing of Environment, 184, 288 – 301, doi: https://doi.org/10.1016/j.rse.2016.07.006, URL http://www.sciencedirect.com/science/article/pii/S0034425716302632, 2016.

Hilley, G. E. and Strecker, M. R.: Processes of oscillatory basin filling and excavation in a tectonically active orogen: Quebrada del Toro Basin, NW Argentina, Geological Society of America Bulletin, 117, 887–901, doi: 10.1130/B25602.1, 2005.

Hofton, M., Dubayah, R., Blair, J. B., and Rabine, D.: Validation of SRTM Elevations Over Vegetated and Nonvegetated Terrain Using Medium Footprint Lidar, Photogrammetric Engineering & Remote Sensing, 72, 279–285, doi: 10.14358/pers.72.3.279, 2006.

Höhle, J. and Höhle, M.: Accuracy assessment of digital elevation models by means of robust statistical methods, ISPRS Journal of Photogrammetry and Remote Sensing, 64, 398–406, doi: 10.1016/j.isprsjprs.2009.02.003, 2009.

Hovius, N., Stark, C. P., and Allen, P. A.: Sediment flux from a mountain belt derived by landslide mapping, Geology, 25, 231–234, 1997.

Huete, A., Justice, C., and Liu, H.: Development of vegetation and soil indices for MODIS-EOS, Remote Sensing of Environment, 49, 224–234, 1994.

Hwang, P. A. and Fois, F.: Surface roughness and breaking wave properties retrieved from polarimetric microwave radar backscattering, Journal of Geophysical Research: Oceans, 120, 3640–3657, doi: 10.1002/2015JC010782, URL https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1002/2015JC010782, 2015.

Ibbeken, H. and Schleyer, R.: Photo-sieving: A method for grain-size analysis of coarse-grained, unconsolidated bedding surfaces, Earth Surface Processes and Landforms, 11, 59–77, doi: 10.1002/esp.3290110108, 1986.

Immerzeel, W., Kraaijenbrink, P., Shea, J., Shrestha, A., Pellicciotti, F., Bierkens, M., and De Jong, S.: High-resolution monitoring of Himalayan glacier dynamics using unmanned aerial vehicles, Remote Sensing of Environment, 150, 93–103, doi: 10.1016/j.rse.2014.04.025, 2014.

Jacobs, T. D. B., Junge, T., and Pastewka, L.: Quantitative characterization of surface topography using spectral analysis, Surface Topography: Metrology and Properties, 5, 013 001, doi: 10.1088/2051-672x/aa51f8, URL https://doi.org/10.1088%2F2051-672x%2Faa51f8, 2017.

James, M. R. and Robson, S.: Mitigating systematic error in topographic models derived from UAV and ground-based image networks, Earth Surface Processes and Landforms, 39, 1413–1420, doi: 10.1002/esp.3609, 2014.

Jarvis, A., Reuter, H. I., Nelson, A., Guevara, E., et al.: Hole-filled SRTM for the globe Version 4, available from the CGIAR-CSI SRTM 90m Database (http://srtm. csi. cgiar. org), 2008.

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

Kääb, A.: Remote sensing of mountain glaciers and permafrost creep, vol. 48, Geograph. Inst. d. Univ., 2005.

Kankaku, Y., Suzuki, S., and Osawa, Y.: ALOS-2 mission and development status, in: 2013 IEEE International Geoscience and Remote Sensing Symposium-IGARSS, pp. 2396–2399, IEEE, doi: 10.1109/IGARSS.2013.6723302, 2013.

Kellerhals, R. and Bray, D. I.: Sampling procedures for coarse fluvial sediments, Journal of the Hydraulics Division, 97, 1165–1180, 1971.

Kierein-Young, K.: The integration of optical and radar data to characterize mineralogy and morphology of surfaces in Death Valley, California, U.S.A., International Journal of Remote Sensing, 18, 1517–1541, doi: 10.1080/014311697218250, 1997.

Kondolf, G. M.: PROFILE: hungry water: effects of dams and gravel mining on river channels, Environmental management, 21, 533–551, doi: 10.1007/s002679900048, 1997.

Kondolf, G. M. and Wolman, M. G.: The sizes of salmonid spawning gravels, Water Resources Research, 29, 2275–2285, doi: 10.1029/93WR00402, 1993.

Krieger, G., Moreira, A., Fiedler, H., Hajnsek, I., Werner, M., Younis, M., and Zink, M.: TanDEM-X: A satellite formation for high-resolution SAR interferometry, IEEE Transactions on Geoscience and Remote Sensing, 45, 3317–3341, doi: 10.1109/TGRS.2007.900693, 2007.

Krieger, G., Zink, M., Bachmann, M., Bräutigam, B., Schulze, D., Martone, M., Rizzoli, P., Steinbrecher, U., Walter Antony, J., De Zan, F., Hajnsek, I., Papathanassiou, K., Kugler, F., Rodriguez Cassola, M., Younis, M., Baumgartner, S., López-Dekker, P., Prats, P., and Moreira, A.: TanDEM-X: A radar interferometer with two formation-flying satellites, Acta Astronautica, 89, 83–98, doi: 10.1016/j.actaastro.2013.03.008, 2013.

Lamb, M. P. and Venditti, J. G.: The grain size gap and abrupt gravel-sand transitions in rivers due to suspension fallout, Geophysical Research Letters, 43, 3777–3785, doi: 10.1002/2016GL068713, 2016.

Lanczos, C.: An iteration method for the solution of the eigenvalue problem of linear differential and integral operators, J. Res. Natl. Bur. Stand. B, 45, 255–282, 1950.

Lane, S. N., Westaway, R. M., and Murray Hicks, D.: Estimation of erosion and deposition volumes in a large, gravel-bed, braided river using synoptic remote sensing, Earth Surface Processes and Landforms, 28, 249–271, 2003.

Langhammer, J., Lendzioch, T., Miřijovský, J., and Hartvich, F.: UAV-Based Optical Granulometry as Tool for Detecting Changes in Structure of Flood Depositions, Remote Sensing, 9, doi: 10.3390/rs9030240, 2017.

Lloyd, S.: Least squares quantization in PCM, IEEE Transactions on Information Theory, 28, 129–137, doi: 10.1109/TIT.1982.1056489, 1982.

Lucieer, A., Turner, D., King, D. H., and Robinson, S. A.: Using an Unmanned Aerial Vehicle (UAV) to capture micro-topography of Antarctic moss beds, International Journal of Applied Earth Observation and Geoinformation, 27, 53–62, doi: 10.1016/j.jag.2013.05.011, 2014.

Marrett, R., Allmendinger, R., Alonso, R., and Drake, R.: Late Cenozoic tectonic evolution of the Puna Plateau and adjacent foreland, northwestern Argentine Andes, Journal of South American Earth Sciences, 7, 179–207, 1994.

Mason, J. and Mohrig, D.: Using Time-Lapse Lidar to Quantify River Bend Evolution on the Meandering Coastal Trinity River, Texas, USA, Journal of Geophysical Research: Earth Surface, 123, doi: 10.1029/2017JF004492, 2018.

McCauley, J. F., Schaber, G. G., Breed, C. S., Grolier, M. J., Haynes, C. V., Issawi, B., Elachi, C., and Blom, R.: Subsurface Valleys and Geoarcheology of the Eastern Sahara Revealed by Shuttle Radar, Science, 218, 1004–1020, doi: 10.1126/science.218.4576.1004, URL https://science.sciencemag.org/content/218/4576/1004, 1982.

Mushkin, A. and Gillespie, A.: Estimating sub-pixel surface roughness using remotely sensed stereoscopic data, Remote Sensing of Environment, 99, 75 – 83, doi: https://doi.org/10.1016/j.rse.2005.02.018, URL http://www.sciencedirect.com/science/article/pii/S003442570500177X, scientific Results from ASTER, 2005.

Mushkin, A. and Gillespie, A. R.: Mapping sub-pixel surface roughness on Mars using high-resolution satellite image data, Geophysical Research Letters, 33, doi: 10.1029/2006GL027095, URL https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2006GL027095, 2006.

Neelmeijer, J., Motagh, M., and Bookhagen, B.: High-resolution digital elevation models from single-pass TanDEM-X interferometry over mountainous regions: A case study of Inylchek Glacier, Central Asia, ISPRS Journal of Photogrammetry and Remote Sensing, 130, 108–121, doi: 10.1016/j.isprsjprs.2017.05.011, 2017.

Nuimura, T., Fujita, K., Yamaguchi, S., and Sharma, R. R.: Elevation changes of glaciers revealed by multitemporal digital elevation models calibrated by GPS survey in the Khumbu region, Nepal Himalaya, 1992–2008, Journal of Glaciology, 58, 648–656, 2012.

Nuth, C. and Kääb, A.: Co-registration and bias corrections of satellite elevation data sets for quantifying glacier thickness change, The Cryosphere, 5, 271–290, doi: 10.5194/tc-5-271-2011, 2011.

Olen, S. and Bookhagen, B.: Mapping damage-affected areas after natural hazard events using Sentinel-1 coherence time series, Remote Sensing, 10, 1272, doi: 10.3390/rs10081272, 2018.

Olen, S. M., Bookhagen, B., and Strecker, M. R.: Role of climate and vegetation density in modulating denudation rates in the Himalaya, Earth and Planetary Science Letters, 445, 57–67, doi: 10.1016/j.epsl.2016.03.047, 2016.

Oskin, M. E., Arrowsmith, J. R., Corona, A. H., Elliott, A. J., Fletcher, J. M., Fielding, E. J., Gold, P. O., Garcia, J. J. G., Hudnut, K. W., Liu-Zeng, J., et al.: Near-field deformation from the El Mayor–Cucapah earthquake revealed by differential LIDAR, Science, 335, 702–705, 2012.

Otsu, N.: A Threshold Selection Method from Gray-Level Histograms, IEEE Transactions on Systems, Man, and Cybernetics, 9, 62–66, doi: 10.1109/TSMC.1979.4310076, 1979.

Ouimet, W. B., Whipple, K. X., and Granger, D. E.: Beyond threshold hillslopes: Channel adjustment to base-level fall in tectonically active mountain ranges, Geology, 37, 579–582, doi: 10.1130/G30013A.1, 2009.

Paola, C., Parker, G., Seal, R., Sinha, S. K., Southard, J. B., and Wilcock, P. R.: Downstream Fining by Selective Deposition in a Laboratory Flume, Science, 258, 1757–1760, doi: 10.1126/science.258.5089.1757, 1992.

Parker, G., Klingeman, P. C., and McLean, D. G.: Bedload and size distribution in paved gravel-bed streams, Journal of the Hydraulics Division, 108, 544–571, 1982.

Passalacqua, P., Belmont, P., Staley, D. M., Simley, J. D., Arrowsmith, J. R., Bode, C. A., Crosby, C., DeLong, S. B., Glenn, N. F., Kelly, S. A., Lague, D., Sangireddy, H., Schaffrath, K., Tarboton, D. G., Wasklewicz, T., and Wheaton, J. M.: Analyzing high resolution topography for advancing the understanding of mass and energy transfer through landscapes: A review, Earth-Science Reviews, 148, 174–193, doi: 10.1016/j.earscirev.2015.05.012, 2015.

Paul, F.: Calculation of glacier elevation changes with SRTM: is there an elevation-dependent bias?, Journal of Glaciology, 54, 945–946, 2008.

Peake, W. and Oliver, T.: The response of terrestrial surfaces at microwave frequencies, Tech. Rep. AFAL-TR-70-301, U.S. Air Force Avionics Lab, 1971.

Pearson, E., Smith, M., Klaar, M., and Brown, L.: Can high resolution 3D topographic surveys provide reliable grain size estimates in gravel bed rivers?, Geomorphology, 293, 143–155, doi: 10.1016/j.geomorph.2017.05.015, 2017.

Peduzzi, P., Herold, C., and Silverio Torres, W. C.: Assessing high altitude glacier thickness, volume and area changes using field, GIS and remote sensing techniques: the case of Nevado Coropuna (Peru), Cryosphere, 4, 313–323, 2010.

Perron, J. T., Kirchner, J. W., and Dietrich, W. E.: Spectral signatures of characteristic spatial scales and nonfractal structure in landscapes, Journal of Geophysical Research: Earth Surface, 113, 2008.

Perroy, R. L., Bookhagen, B., Asner, G. P., and Chadwick, O. A.: Comparison of gully erosion estimates using airborne and ground-based LiDAR on Santa Cruz Island, California, Geomorphology, 118, 288–300, 2010.

Pfeiffer, A. M., Finnegan, N. J., and Willenbring, J. K.: Sediment supply controls equilibrium channel geometry in gravel rivers, Proceedings of the National Academy of Sciences, 114, 3346–3351, 2017.

Pike, R., Evans, I., and Hengl, T.: Chapter 1 Geomorphometry: A Brief Guide, in: Developments in Soil Science, vol. 33, pp. 3–30, Elsevier BV, doi: 10.1016/S0166-2481(08)00001-9, 2009.

Pingel, H., Strecker, M. R., Alonso, R. N., and Schmitt, A. K.: Neotectonic basin and landscape evolution in the Eastern Cordillera of NW Argentina, Humahuaca Basin (˜ 24 S), Basin Research, 25, 554–573, doi: 10.1111/bre.12016, 2013.

Purinton, B. and Bookhagen, B.: Validation of digital elevation models (DEMs) and comparison of geomorphic metrics on the southern Central Andean Plateau, Earth Surface Dynamics, 5, 211–237, doi: 10.5194/esurf-5-211-2017, URL https://www.earth-surf-dynam.net/5/211/2017/, 2017.

Purinton, B. and Bookhagen, B.: Measuring decadal vertical land-level changes from SRTM-C (2000) and TanDEM-X ($\sim$ 2015) in the south-central Andes, Earth Surface Dynamics, 6, 971–987, doi: 10.5194/esurf-6-971-2018, 2018.

Purinton, B. and Bookhagen, B.: PebbleCounts: a Python grain-sizing algorithm for gravel-bed river imagery, doi: 10.5880/fidgeo.2019.007, URL https://github.com/UP-RS-ESP/PebbleCounts, 2019a.

Purinton, B. and Bookhagen, B.: Introducing *PebbleCounts*: a grain-sizing tool for photo surveys of dynamic gravel-bed rivers, Earth Surface Dynamics, 7, 859–877, doi: 10.5194/esurf-7-859-2019, URL https://www.earth-surf-dynam.net/7/859/2019/, 2019b.

Racoviteanu, A. E., Manley, W. F., Arnaud, Y., and Williams, M. W.: Evaluating digital elevation models for glaciologic

applications: An example from Nevado Coropuna, Peruvian Andes, Global and Planetary Change, 59, 110–125, doi: 10.1016/j.gloplacha.2006.11.036, 2007.

Rahman, M., Moran, M., Thoma, D., Bryant, R., Collins, C. H., Jackson, T., Orr, B., and Tischler, M.: Mapping surface roughness and soil moisture using multi-angle radar imagery without ancillary data, Remote Sensing of Environment, 112, 391 – 402, doi: https://doi.org/10.1016/j.rse.2006.10.026, URL http://www.sciencedirect.com/science/article/pii/S0034425707003008, soil Moisture Experiments 2004 (SMEX04) Special Issue, 2008.

Regmi, P., Grosse, G., Jones, M. C., Jones, B. M., and Anthony, K. W.: Characterizing Post-Drainage Succession in Thermokarst Lake Basins on the Seward Peninsula, Alaska with TerraSAR-X Backscatter and Landsat-based NDVI Data, Remote Sensing, 4, 3741–3765, doi: 10.3390/rs4123741, URL https://www.mdpi.com/2072-4292/4/12/3741, 2012.

Rexer, M. and Hirt, C.: Comparison of free high resolution digital elevation data sets (ASTER GDEM2, SRTM v2.1/v4.1) and validation against accurate heights from the Australian National Gravity Database, Australian Journal of Earth Sciences, 61, 213–226, doi: 10.1080/08120099.2014.884983, 2014.

Rheinwalt, A., Goswami, B., and Bookhagen, B.: A Network-Based Flow Accumulation Algorithm for Point Clouds: Facet-Flow Networks (FFNs), Journal of Geophysical Research: Earth Surface, 124, 2013–2033, doi: 10.1029/2018JF004827, 2019.

Rice, S. and Church, M.: Sampling surficial fluvial gravels; the precision of size distribution percentile sediments, Journal of Sedimentary Research, 66, 654, doi: 10.2110/jsr.66.654, 1996.

Rice, S. and Church, M.: Grain size along two gravel-bed rivers: statistical variation, spatial pattern and sedimentary links, Earth Surface Processes and Landforms, 23, 345–363, doi: 10.1002/(SICI)1096-9837(199804)23:4<345::AID-ESP850>3.0.CO;2-B, 1998.

Ridley, J., Strawbridge, F., Card, R., and Phillips, H.: Radar backscatter characteristics of a desert surface, Remote Sensing of Environment, 57, 63 – 78, doi: https://doi.org/10.1016/0034-4257(96)00018-1, URL http://www.sciencedirect.com/science/article/pii/0034425796000181, 1996.

Rignot, E., Echelmeyer, K., and Krabill, W.: Penetration depth of interferometric synthetic-aperture radar signals in snow and ice, Geophysical Research Letters, 28, 3501–3504, doi: 10.1029/2000GL012484, 2001.

Rinaldi, M. and Simon, A.: Bed-level adjustments in the Arno River, central Italy, Geomorphology, 22, 57–71, 1998.

Rizzoli, P., Martone, M., Gonzalez, C., Wecklich, C., Tridon, D. B., Bräutigam, B., Bachmann, M., Schulze, D., Fritz, T., Huber, M., et al.: Generation and performance assessment of the global TanDEM-X digital elevation model, ISPRS Journal of Photogrammetry and Remote Sensing, 132, 119–139, 2017.

Rodríguez, E., Morris, C. S., and Belz, J. E.: A Global Assessment of the SRTM Performance, Photogrammetric Engineering & Remote Sensing, 72, 249–260, doi: 10.14358/pers.72.3.249, 2006.

Roering, J., Perron, J., and Kirchner, J.: Hillslope morphology and functional relationships between topographic relief and denudation, Earth and Planetary Science Letters, 264, 245–258, doi: 10.1016/j.epsl.2007.09.035, 2007.

Roering, J. J., Kirchner, J. W., and Dietrich, W. E.: Evidence for nonlinear, diffusive sediment transport on hillslopes and implications for landscape morphology, Water Resources Research, 35, 853–870, doi: 10.1029/1998WR900090, 1999.

Roering, J. J., Mackey, B. H., Marshall, J. A., Sweeney, K. E., Deligne, N. I., Booth, A. M., Handwerger, A. L., and Cerovski-Darriau, C.: 'You are HERE': Connecting the dots with airborne lidar for geomorphic fieldwork, Geomorphology, 200, 172–183, doi: 10.1016/j.geomorph.2013.04.009, 2013.

Rossi, C., Minet, C., Fritz, T., Eineder, M., and Bamler, R.: Temporal monitoring of subglacial volcanoes with TanDEM-X — Application to the 2014–2015 eruption within the Bárðarbunga volcanic system, Iceland, Remote Sensing of Environment, 181, 186–197, doi: 10.1016/j.rse.2016.04.003, 2016.

Rossini, M., Di Mauro, B., Garzonio, R., Baccolo, G., Cavallini, G., Mattavelli, M., De Amicis, M., and Colombo, R.: Rapid melting dynamics of an alpine glacier with repeated UAV photogrammetry, Geomorphology, 304, 159–172, doi: 10.1016/j.geomorph.2017.12.039, 2018.

Rovira, A., Batalla, R., and Sala, M.: Response of a river sediment budget after historical gravel mining (the lower Tordera, NE Spain), River Research and Applications, 21, 829–847, 2005.

Rubin, D. M.: A Simple Autocorrelation Algorithm for Determining Grain Size from Digital Images of Sediment, Journal of Sedimentary Research, 74, 160, doi: 10.1306/052203740160, 2004.

Ruiz, L., Berthier, E., Viale, M., Pitte, P., and Masiokas, M. H.: Recent geodetic mass balance of Monte Tronador glaciers, northern Patagonian Andes, The Cryosphere, 11, 619, 2017.

Russ, J. C.: The image processing handbook, fourth edition, CRC press, 2002.

Rychkov, I., Brasington, J., and Vericat, D.: Computational and methodological aspects of terrestrial surface analysis based on point clouds, Computers & Geosciences, 42, 64–70, doi: 10.1016/j.cageo.2012.02.011, 2012.

Sadeh, Y., Cohen, H., Maman, S., and Blumberg, D. G.: Evaluation of Manning's n Roughness Coefficient in Arid Environments by Using SAR Backscatter, Remote Sensing, 10, doi: 10.3390/rs10101505, URL https://www.mdpi.com/2072-4292/10/10/1505, 2018.

Sangireddy, H., Stark, C. P., and Passalacqua, P.: Multiresolution analysis of characteristic length scales with high-resolution topographic data, Journal of Geophysical Research: Earth Surface, 122, 1296–1324, doi: 10.1002/2015JF003788, 2015JF003788, 2017.

Savi, S., Schildgen, T. F., Tofelde, S., Wittmann, H., Scherler, D., Mey, J., Alonso, R. N., and Strecker, M. R.: Climatic controls on debris-flow activity and sediment aggradation: The Del Medio fan, NW Argentina, Journal of Geophysical Research: Earth Surface, 121, 2424–2445, 2016.

Scherler, D., Bookhagen, B., Wulf, H., Preusser, F., and Strecker, M. R.: Increased late Pleistocene erosion rates during fluvial aggradation in the Garhwal Himalaya, northern India, Earth and Planetary Science Letters, 428, 255–266, doi: 10.1016/j.epsl.2015.06.034, 2015.

Schildgen, T. F., Robinson, R. A., Savi, S., Phillips, W. M., Spencer, J. Q., Bookhagen, B., Scherler, D., Tofelde, S., Alonso, R. N., Kubik, P. W., et al.: Landscape response to late Pleistocene climate change in NW Argentina: Sediment flux modulated by basin geometry and connectivity, Journal of Geophysical Research: Earth Surface, 121, 392–414, 2016.

Schumm, S. A.: Evolution of drainage systems and slopes in badlands at Perth Amboy, New Jersey, GSA Bulletin, 67, 597–646, doi: 10.1130/0016-7606(1956)67[597:EODSAS]2.0.CO;2, 1956.

Schutz, B., Zwally, H., Shuman, C., Hancock, D., and DiMarzio, J.: Overview of the ICESat mission, Geophysical Research Letters, 32, 2005.

Sculley, D.: Web-scale K-means Clustering, in: Proceedings of the 19th International Conference on World Wide Web, pp. 1177–1178, ACM, New York, NY, USA, doi: 10.1145/1772690.1772862, 2010.

Shean, D. E., Alexandrov, O., Moratto, Z. M., Smith, B. E., Joughin, I. R., Porter, C., and Morin, P.: An automated, open-source pipeline for mass production of digital elevation models (DEMs) from very-high-resolution commercial stereo satellite imagery, ISPRS Journal of Photogrammetry and Remote Sensing, 116, 101–117, doi: 10.1016/j.isprsjprs.2016.03.012, 2016.

Shepard, M. K., Campbell, B. A., Bulmer, M. H., Farr, T. G., Gaddis, L. R., and Plaut, J. J.: The roughness of natural terrain: A planetary and remote sensing perspective, Journal of Geophysical Research: Planets, 106, 32 777–32 795, doi: 10.1029/2000JE001429, URL https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2000JE001429, 2001.

Shields, A.: Anwendung der Aehnlichkeitsmechanik und der Turbulenzforschung auf die Geschiebebewegung, Ph.D. thesis, Technical University Berlin, 1936.

Shortridge, A. and Messina, J.: Spatial structure and landscape associations of SRTM error, Remote Sensing of Environment, 115, 1576–1587, doi: 10.1016/j.rse.2011.02.017, 2011.

Sime, L. and Ferguson, R.: Information on Grain Sizes in Gravel-Bed Rivers by Automated Image Analysis, Journal of Sedimentary Research, 73, 630, doi: 10.1306/112102730630, 2003.

Sklar, L. S., Dietrich, W. E., Foufoula-Georgiou, E., Lashermes, B., and Bellugi, D.: Do gravel bed river size distributions record channel network structure?, Water Resources Research, 42, W06D18, doi: 10.1029/2006WR005035, 2006.

Small, D.: Flattening gamma: Radiometric terrain correction for SAR imagery, IEEE Transactions on Geoscience and Remote Sensing, 49, 3081–3093, doi: 10.1109/TGRS.2011.2120616, 2011.

Smith, M., Carrivick, J., and Quincey, D.: Structure from motion photogrammetry in physical geography, Progress in Physical Geography: Earth and Environment, 40, 247–275, doi: 10.1177/0309133315615805, 2016.

Smith, T., Rheinwalt, A., and Bookhagen, B.: Determining the optimal grid resolution for topographic analysis on an airborne lidar dataset, Earth Surface Dynamics, 7, 475–489, doi: 10.5194/esurf-7-475-2019, 2019.

SNAP, E.: SNAP Software, URL https://step.esa.int/main/toolboxes/snap/, 2019.

Snavely, N., Seitz, S. M., and Szeliski, R.: Modeling the world from internet photo collections, International journal of computer vision, 80, 189–210, doi: 10.1007/s11263-007-0107-3, 2008.

Snyder, N. P., Whipple, K. X., Tucker, G. E., and Merritts, D. J.: Landscape response to tectonic forcing: Digital elevation model analysis of stream profiles in the Mendocino triple junction region, northern California, Geological Society of America Bulletin, 112, 1250–1263, doi: 10.1130/0016-7606(2000)112<1250:lrttfd>2.3.co;2, 2000.

Sobel, E. R., Hilley, G. E., and Strecker, M. R.: Formation of internally drained contractional basins by aridity-limited bedrock incision, Journal of Geophysical Research: Solid Earth, 108, doi: 10.1029/2002JB001883, 2003.

Srivastava, H. S., Patel, P., Sharma, Y., and Navalgund, R. R.: Large-Area Soil Moisture Estimation Using Multi-Incidence-Angle RADARSAT-1 SAR Data, IEEE Transactions on Geoscience and Remote Sensing, 47, 2528–2535, doi: 10.1109/TGRS.2009.2018448, 2009.

Strecker, M., Alonso, R., Bookhagen, B., Carrapa, B., Hilley, G., Sobel, E., and Trauth, M.: Tectonics and climate of the southern central Andes, Annu. Rev. Earth Planet. Sci., 35, 747–787, doi: 10.1146/annurev.earth.35.031306.140158, 2007.

Sun, G., Ranson, K. J., Kharuk, V. I., and Kovacs, K.: Validation of surface height from shuttle radar topography mission using shuttle laser altimeter, Remote Sensing of Environment, 88, 401–411, 2003.

Sundborg, Å.: The river Klarälven a study of fluvial processes, Geografiska annaler, 38, 125–316, 1956.

Syvitski, J. P., Vörösmarty, C. J., Kettner, A. J., and Green, P.: Impact of humans on the flux of terrestrial sediment to the global coastal ocean, Science, 308, 376–380, 2005.

Tachikawa, T., Kaku, M., Iwasaki, A., Gesch, D. B., Oimoen, M. J., Zhang, Z., Danielson, J. J., Krieger, T., Curtis, B., Haase, J., et al.: ASTER global digital elevation model version 2-summary of validation results, Tech. rep., NASA, 2011.

Tadono, T., Ishida, H., Oda, F., Naito, S., Minakawa, K., and Iwamoto, H.: Precise Global DEM Generation by ALOS PRISM, ISPRS Annals of Photogrammetry, Remote Sensing and Spatial Information Sciences, II-4, 71–76, doi: 10.5194/isprsannals-ii-4-71-2014, 2014.

Tarboton, D. G.: A new method for the determination of flow directions and upslope areas in grid digital elevation models, Water Resources Research, 33, 309–319, doi: 10.1029/96WR03137, 1997.

Tarolli, P.: High-resolution topography for understanding Earth surface processes: Opportunities and challenges, Geomorphology, 216, 295–312, doi: 10.1016/j.geomorph.2014.03.008, 2014.

Telea, A.: An Image Inpainting Technique Based on the Fast Marching Method, J. Graphics, GPU, and Game Tools, 9, 23–34, doi: 10.1080/10867651.2004.10487596, 2004.

Tofelde, S., Schildgen, T. F., Savi, S., Pingel, H., Wickert, A. D., Bookhagen, B., Wittmann, H., Alonso, R. N., Cottle, J., and Strecker, M. R.: 100 kyr fluvial cut-and-fill terrace cycles since the Middle Pleistocene in the southern Central Andes, NW Argentina, Earth and Planetary Science Letters, 473, 141–153, doi: 10.1016/j.epsl.2017.06.001, 2017.

Tomasi, C. and Manduchi, R.: Bilateral filtering for gray and color images, in: Sixth International Conference on Computer Vision (IEEE Cat. No.98CH36271), pp. 839–846, doi: 10.1109/ICCV.1998.710815, 1998.

Torres, R., Snoeij, P., Geudtner, D., Bibby, D., Davidson, M., Attema, E., Potin, P., Rommen, B., Floury, N., Brown, M., Traver, I. N., Deghaye, P., Duesmann, B., Rosich, B., Miranda, N., Bruno, C., L'Abbate, M., Croci, R., Pietropaolo, A., Huchler, M., and Rostan, F.: GMES Sentinel-1 mission, Remote Sensing of Environment, 120, 9 – 24, doi: https://doi.org/10.1016/j.rse.2011.05.028, URL http://www.sciencedirect.com/science/article/pii/S0034425712000600, the Sentinel Missions - New Opportunities for Science, 2012.

Tucker, G., Lancaster, S., Gasparini, N., and Bras, R.: The Channel-Hillslope Integrated Landscape Development Model (CHILD), chap. 12, pp. 349–388, Springer US, doi: 10.1007/978-1-4615-0575-4_12, 2001.

Uddin, K., Matin, M. A., and Meyer, F. J.: Operational Flood Mapping Using Multi-Temporal Sentinel-1 SAR Images: A Case Study from Bangladesh, Remote Sensing, 11, doi: 10.3390/rs11131581, URL https://www.mdpi.com/2072-4292/11/13/1581, 2019.

Ulaby, F. T., Moore, R. K., and Fung, A. K.: Microwave Remote Sensing, Volumes 1, 2, and 3, p. 2162, Addison-Wesley, 1982.

van Iersel, W., Straatsma, M., Middelkoop, H., and Addink, E.: Multitemporal classification of river floodplain vegetation using time series of UAV images, Remote Sensing, 10, 1144, doi: 10.3390/rs10071144, 2018.

Van Niel, T. G., McVicar, T. R., Li, L., Gallant, J. C., and Yang, Q.: The impact of misregistration on SRTM and DEM image differences, Remote Sensing of Environment, 112, 2430–2442, doi: 10.1016/j.rse.2007.11.003, 2008.

Vera, C., Higgins, W., Amador, J., Ambrizzi, T., Garreaud, R., Gochis, D., Gutzler, D., Lettenmaier, D., Marengo, J., Mechoso, C., et al.: Toward a unified view of the American monsoon systems, Journal of climate, 19, 4977–5000, 2006.

Verdú, J. M., Batalla, R. J., and Martínez-Casasnovas, J. A.: High-resolution grain-size characterisation of gravel bars using imagery analysis and geo-statistics, Geomorphology, 72, 73–93, doi: 10.1016/j.geomorph.2005.04.015, 2005.

Verhoest, N. E., Lievens, H., Wagner, W., Álvarez Mozos, J., Moran, M. S., and Mattia, F.: On the Soil Roughness Parameterization Problem in Soil Moisture Retrieval of Bare Surfaces from Synthetic Aperture Radar, Sensors, 8, 4213–4248, doi: 10.3390/s8074213, URL https://www.mdpi.com/1424-8220/8/7/4213, 2008.

Verma, A. K. and Bourke, M. C.: A method based on structure-from-motion photogrammetry to generate sub-millimetre-resolution digital elevation models for investigating rock breakdown features, Earth Surface Dynamics, 7, 45–66, doi: 10.5194/esurf-7-45-2019, URL https://www.earth-surf-dynam.net/7/45/2019/, 2019.

Vreugdenhil, M., Wagner, W., Bauer-Marschallinger, B., Pfeil, I., Teubner, I., Rüdiger, C., and Strauss, P.: Sensitivity of Sentinel-1 Backscatter to Vegetation Dynamics: An Austrian Case Study, Remote Sensing, 10, doi: 10.3390/rs10091396, URL https://www.mdpi.com/2072-4292/10/9/1396, 2018.

Walter, R. C. and Merritts, D. J.: Natural streams and the legacy of water-powered mills, Science, 319, 299–304, 2008.

Wang, D. and Kääb, A.: Modeling glacier elevation change from DEM time series, Remote Sensing, 7, 10 117–10 142, 2015.

Ward, J. H.: Hierarchical Grouping to Optimize an Objective Function, Journal of the American Statistical Association, 58, 236–244, doi: 10.1080/01621459.1963.10500845, 1963.

Warrick, J. A., Rubin, D. M., Ruggiero, P., Harney, J. N., Draut, A. E., and Buscombe, D.: Cobble cam: grain-size measurements of sand to boulder from digital photographs and autocorrelation analyses, Earth Surface Processes and Landforms, 34, 1811–1821, doi: 10.1002/esp.1877, 2009.

Weeks, R., Smith, M., Pak, K., and Gillespie, A.: Inversions of SIR-C and AIRSAR data for the roughness of geological surfaces, Remote Sensing of Environment, 59, 383 – 396, doi: https://doi.org/10.1016/S0034-4257(96)00161-7, URL http://www.sciencedirect.com/science/article/pii/S0034425796001617, spaceborne Imaging Radar Mission, 1997.

Weeks, R. J., Smith, M., Pak, K., Li, W.-H., Gillespie, A., and Gustafson, B.: Surface roughness, radar backscatter, and visible and near-infrared reflectance in Death Valley, California, Journal of Geophysical Research: Planets, 101, 23 077–23 090, doi: 10.1029/96JE01247, URL https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/96JE01247, 1996.

Wendleder, A., Wessel, B., Roth, A., Breunig, M., Martin, K., and Wagenbrenner, S.: TanDEM-X water indication mask: Generation and first evaluation results, IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 6, 171–179, 2013.

Wessel, B.: TanDEM-X Ground Segment–DEM Products Specification Document, Technical Note 3.1, German Aerospace Center (DLR) (Aug.)., 2016.

Wessel, B., Huber, M., Wohlfart, C., Marschalk, U., Kosmann, D., and Roth, A.: Accuracy Assessment of the Global TanDEM-X Digital Elevation Model with GPS Data, ISPRS Journal of Photogrammetry and Remote Sensing, pp. 1–12, 2018.

Westoby, M. J., Brasington, J., Glasser, N. F., Hambrey, M. J., and Reynolds, J. M.: 'Structure-from-Motion'photogrammetry: A low-cost, effective tool for geoscience applications, Geomorphology, 179, 300–314, doi: 10.1016/j.geomorph.2012.08.021, 2012.

Westoby, M. J., Dunning, S. A., Woodward, J., Hein, A. S., Marrero, S. M., Winter, K., and Sugden, D. E.: Sedimentological characterization of Antarctic moraines using UAVs and Structure-from-Motion photogrammetry, Journal of

Glaciology, 61, 1088–1102, doi: 10.3189/2015JoG15J086, 2015.

Wheaton, J. M., Brasington, J., Darby, S. E., and Sear, D. A.: Accounting for uncertainty in DEMs from repeat topographic surveys: improved sediment budgets, Earth Surface Processes and Landforms, 35, 136–156, 2010.

Williams, K. K. and Greeley, R.: Laboratory and field measurements of the modification of radar backscatter by sand, Remote Sensing of Environment, 89, 29 – 40, doi: https://doi.org/10.1016/j.rse.2003.09.006, URL http://www.sciencedirect.com/science/article/pii/S003442570300258X, 2004.

Wohl, E. E., Anthony, D. J., Madsen, S. W., and Thompson, D. M.: A comparison of surface sampling methods for coarse fluvial sediments, Water Resources Research, 32, 3219–3226, doi: 10.1029/96WR01527, 1996.

Wolcott, J. and Church, M.: Strategies for sampling spatially heterogeneous phenomena; the example of river gravels, Journal of Sedimentary Research, 61, 534–543, doi: 10.1306/D4267753-2B26-11D7-8648000102C1865D, 1991.

Wolman, M. G.: A method of sampling coarse river-bed material, Eos, Transactions American Geophysical Union, 35, 951–956, doi: 10.1029/TR035i006p00951, 1954.

Woodget, A. S. and Austrums, R.: Subaerial gravel size measurement using topographic data derived from a UAV-SfM approach, Earth Surface Processes and Landforms, 42, 1434–1443, doi: 10.1002/esp.4139, 2017.

Woodget, A. S., Fyffe, C., and Carbonneau, P. E.: From manned to unmanned aircraft: Adapting airborne particle size mapping methodologies to the characteristics of sUAS and SfM, Earth Surface Processes and Landforms, 43, 857–870, doi: 10.1002/esp.4285, 2018.

Yamazaki, D., Ikeshima, D., Tawatari, R., Yamaguchi, T., O'Loughlin, F., Neal, J. C., Sampson, C. C., Kanae, S., and Bates, P. D.: A high accuracy map of global terrain elevations, Geophysical Research Letters, 44, 5844–5853, doi: 10.1002/2017GL072874, 2017.

Zwally, H., Schutz, R., Bentley, C., Bufton, J., Herring, T., Minster, J., Spinhirne, J., and Thomas, R.: GLAS/ICESat L2 Global Land Surface Altimetry Data, version34, doi: 10.5067/ICESAT/GLAS/DATA227, 2009.

Appendix A

# Supporting Information for Chapter 3

## TanDEM-X and SRTM-C Internal Error Exploration

TanDEM-X tiles are delivered with a number of auxiliary layers allowing deeper exploration of errors. The detailed description of each layer can be found in the product technical document (Wessel, 2016). In general we expect the TanDEM-X to have superior accuracy to previous datasets given the fine raw radiometric resolution (~3.3 m) combined with the high number of contributing scenes per pixel (Fig. A1). Vertical uncertainties from our extensive differential GPS (dGPS) control dataset (Purinton and Bookhagen, 2017) revealed the dependence of vertical uncertainty on both consistency of the input scenes (COM) and the height error map (HEM). The HEM contains expected random error (in meters) for each pixel and is typically larger over less coherent areas like vegetation or water (Wessel et al., 2018). The COM demonstrates the agreement between individual scenes, providing information about pixels which had potential contribution from poor performance TerraSAR-X / TanDEM-X looks. Figures A2–A3 demonstrate these errors compared with sub-20 cm vertical accuracy dGPS points.

The dependence of TanDEM-X error on terrain slope is demonstrated in Figure A4, where we note worsening vertical error with increasing slope and apparent over-prediction of TanDEM-X heights. We see this same trend in the SRTM-C data compared with dGPS (Fig. A5). Since both DEMs are over-predicting heights on steeper slopes in the same way, we expect the relative difference between each to remain the same for the mapping of vertical land-level changes. We note the dearth of height measurements on higher slopes, however, this positive relationship for TanDEM-X is also found by Wessel et al. (2018).

In Figure A6 we explore the relation between HEM and slope including information from the water indication mask (WAM) from TanDEM-X. The WAM is based on amplitude and coherence thresholds (Wendleder et al., 2013) and may therefore contain additional information besides water pixels, such as areas that lost coherence in deserts or very steep and vegetated slopes (Fig. 3.1C). From Figure A6, we note that the very high HEM values at low slopes are reduced when removing pixels identified as having low amplitude or coherence. While these are largely water bodies in TanDEM-X, they also correspond to DEM artifacts on hillslopes and in salt flats. We can observe a non-linear trend with increasing slope and vertical uncertainty, which we also see in the boxplots of Figures A4–A5.

## TanDEM-X and SRTM-C Differencing Errors

In addition to comparison with dGPS and TanDEM-X auxiliary layers, we also examine the dependence on TanDEM-X−SRTM-C *dh* on a number of terrain characteristics. These biases could affect *dh* change mapping and measurement and therefore warrant special attention. Figure A7A demonstrates the expected increase in uncertainty with slope, however, since both the DEMs over-predict heights at higher

slopes in the same manner, we do not observe any linear trends in differencing these datasets at higher slopes. Figure A7B shows that this increase in uncertainty is also related to relief (which increases with higher slopes as well), though local relief has a smaller effect on uncertainty.

Since slope and relief may overprint other biases, we also test the dependence of *dh* on elevation (Fig. A8) and vegetation (Fig. A9) on only low slope and low relief terrain using cutoffs of 5° and 100 m, respectively. Since most high slope and high relief terrain is found in the steep topographic transition zone at the mountain front (Fig. 3.1), this threshold removes all elevations in this band (~1800–2500 m). This orographic mount-front band also corresponds to moderate vegetation, thus only the very low (Altiplano-Puna Plateau) and very high (wet foreland) vegetation values are shown. We note no trends in either vegetation or elevation. Although we would expect vegetation differences between X-band and C-band, it may be the case that sub-tropical forests in the region are not dense enough to fully block radar penetration over a significant number of pixels.

Despite having the same nominal resolution of 1 arcsec, the SRTM-C and TanDEM-X demonstrate clear differences in raw radiometric resolution (~30 m versus ~3.3 m) when comparing plots of terrain curvature versus *dh* (Fig. A10). Many studies of glacier height change apply a polynomial correction to height values based on this relationship (e.g., Gardelle et al., 2012; Ruiz et al., 2017), however, this introduces artificial elevations, whereas we are interested in the raw data. Furthermore, this curvature bias only affects the most peaked and narrowest parts of the landscape, whereas the more planar regions (e.g., gravel-bed channels, landslide lobes) show less error dependence on curvature.

# Relation of Destriped Block Median *dh* to Elevation

Following our non-aggressive ($97.5^{th}$ percentile cutoff and 5% RMSE convergence) FFT destriping, a number of patches of positive and negative *dh* remain. Though these patches were removed in a more aggressive approach ($95^{th}$ percentile cutoff and 2% RMSE convergence), this method also filtered true topography (Fig. A11). We instead rectified these patches using the blocked shifting scheme discussed in Chapter 3. We did not find any clear relation between the medians in each block or elevation for any of the block sizes tested (Fig. A12). Rather the remaining error patterns were likely caused by a complex combination of remaining SRTM-C orbital biases, atmospheric conditions at the time of SRTM-C capture, local adjustment of both datasets to ICESat measurements, and possibly the over-prediction of elevation in both datasets on steep topography (Fig. A4–A5).

# Increase in *dh* Measurements Downstream of Knickpoints

As mentioned in Chapter 3, there is a clear increase in number of measurements downstream of the knickpoint for the Río Grande and Río Toro. The 2D KDE shown in Figure 3.7 does not provide the downstream number of measurements, but is rather a demonstration of relative point density. In Figure A13, we demonstrate the 1 km binned measurements downstream using a simple histogram plot. Here we see the increase in number of measurements in the wider, downstream channel reaches.

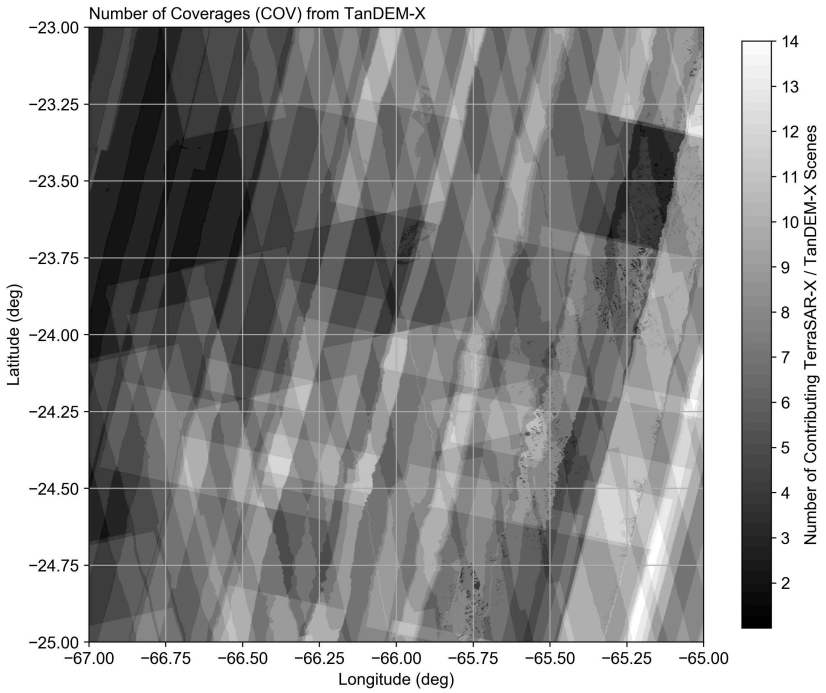


**Figure A1** – Coverage map (COV) for four tiles of TanDEM-X (S 24–25°, W 66–67°). Note high number of coverages up to 14 on eastern side where topography is steeper, versus less coverage to west on Altiplano-Puna low-relief plateau.

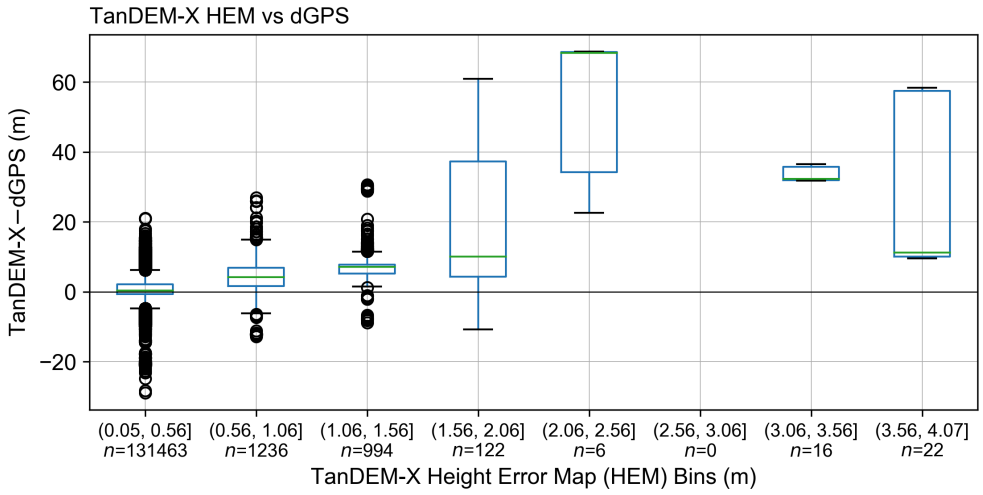


**Figure A2** – Height error map (HEM) values binned in 0.5 m intervals with boxplots showing TanDEM-X elevation compared with dGPS measurements. Note number of measurements ($n$) at higher HEM values drops significantly. Similar to Wessel et al. (2018) we observe a positive trend in vertical uncertainty with HEM and much worse accuracy at HEM values > 0.5 m. These higher HEM values also correspond to higher slopes, indicating the over-prediction of TanDEM-X at extreme slopes (cf. Fig. A4).

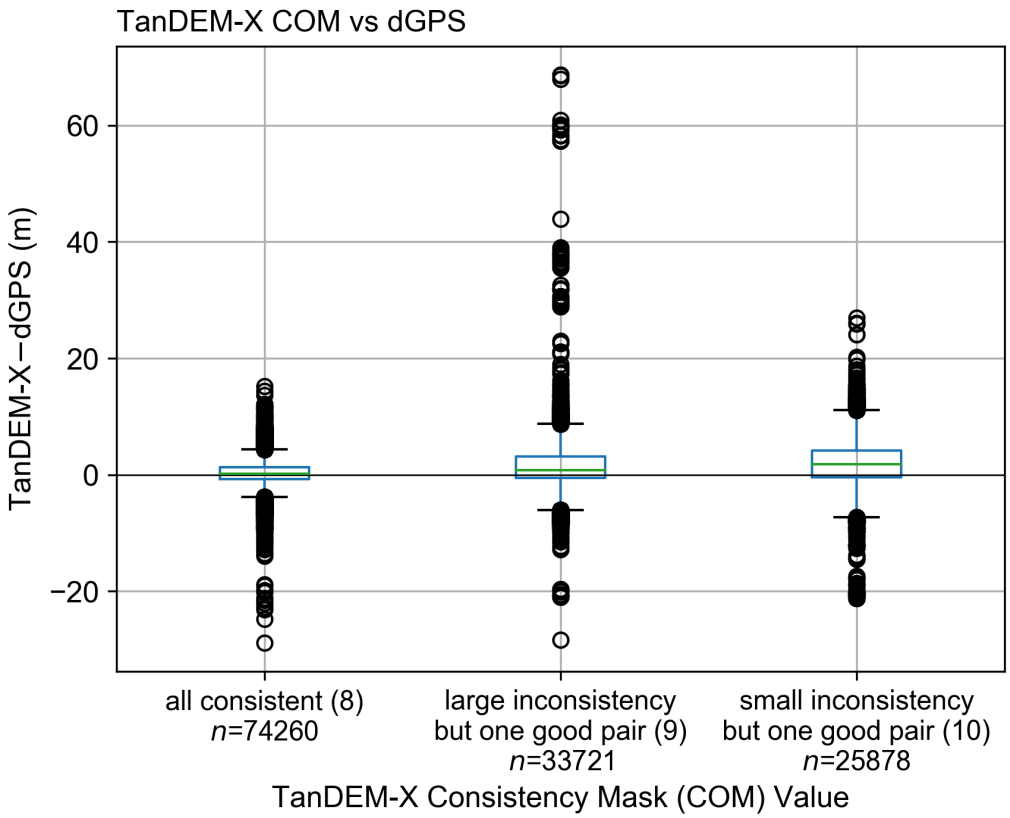


**Figure A3** – Consistency map (COM) boxplots of TanDEM-X vertical uncertainty versus dGPS for all COM values covered, and the meaning of each COM value. See Wessel (2016) for detailed description of each COM value. We note a greater vertical uncertainty with more inconsistent COM values, whereas for COM value 8 (all input heights consistent), the vertical uncertainty boxplot is narrow and centered on zero median.

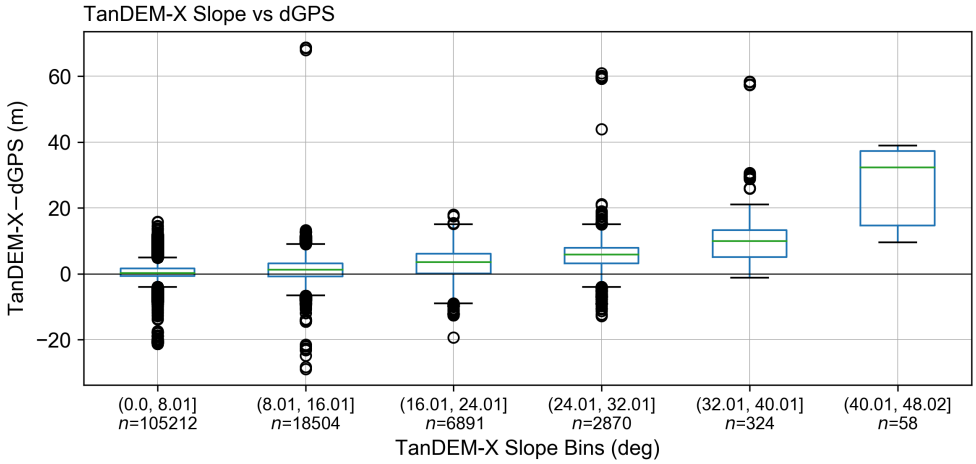


**Figure A4** – Terrain slope binned in 8° intervals versus boxplots of dGPS vertical uncertainty for TanDEM-X. We note worsening vertical uncertainty at higher slopes and an apparent over-prediction of heights by TanDEM-X. Note the drop in number of measurements ($n$) at higher slopes.

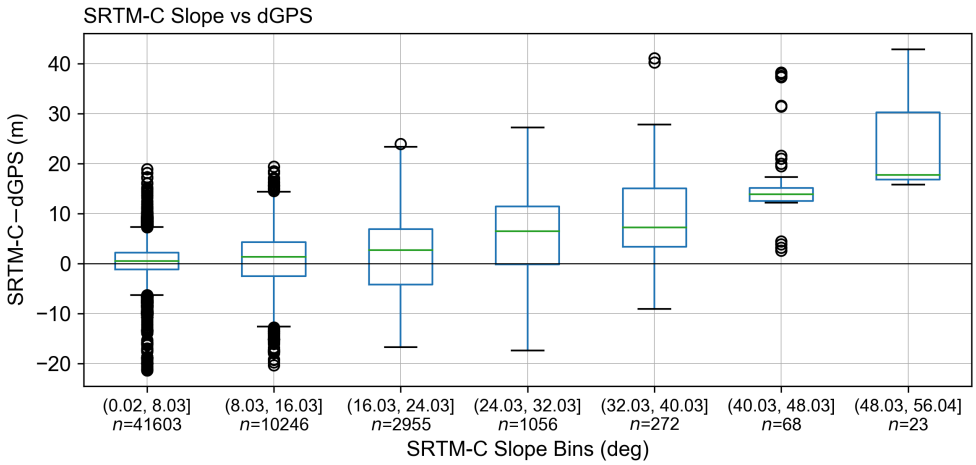


**Figure A5** – Terrain slope binned in 8° intervals versus boxplots of dGPS vertical uncertainty for SRTM-C. We note worsening vertical uncertainty at higher slopes and an apparent over-prediction of heights by SRTM-C. Note the drop in number of measurements ($n$) at higher slopes.

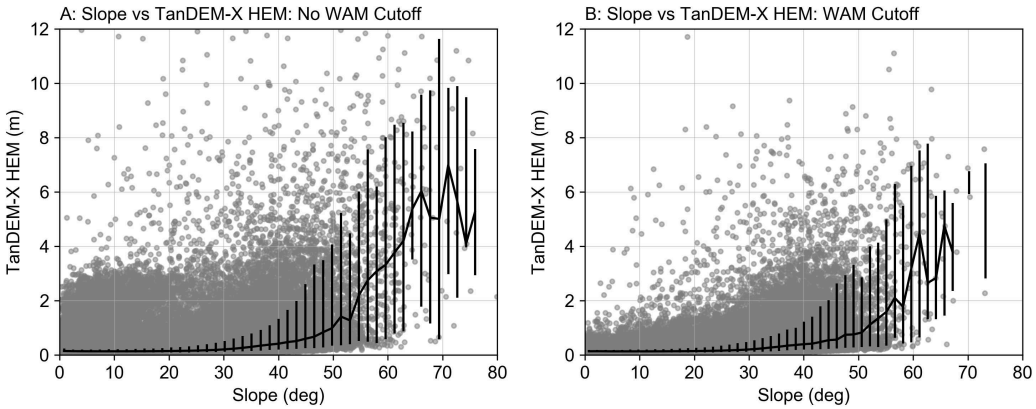


**Figure A6** – Terrain slope versus HEM for TanDEM-X. Gray points are raw data and black line is median values with $10^{th}$ and $90^{th}$ percentile error bars. (A) is all pixels including those with low amplitude and coherence, whereas in (B) the WAM cutoff is applied to remove these pixels. Note the reduction of high HEM outliers at primarily lower slopes (water bodies) with WAM cutoff.

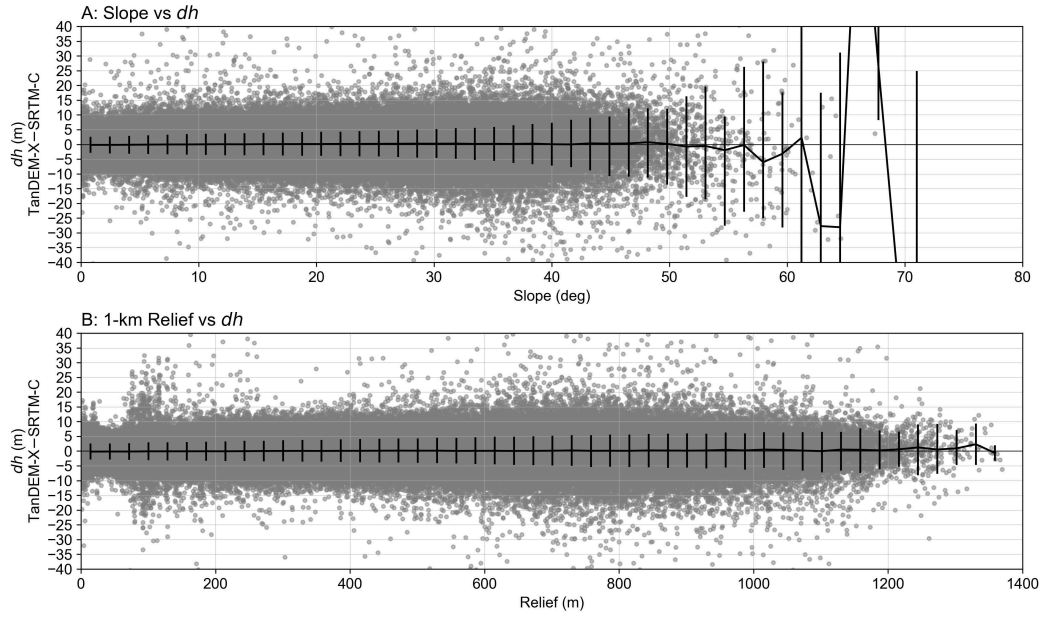


**Figure A7** – (A) Terrain slope versus *dh* and (B) 1-km relief versus *dh*. Gray points are raw data and black line is median values with $10^{th}$ and $90^{th}$ percentile error bars. Aside from greater uncertainties at higher slopes, we note no linear trend with either parameter that may affect *dh* identification.

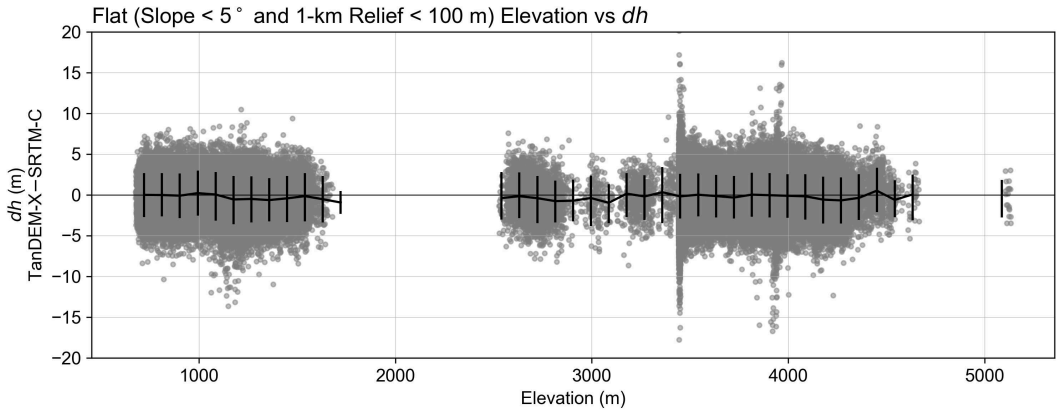


**Figure A8** – Terrain elevation versus *dh*. Gray points are raw data and black line is median values with $10^{th}$ and $90^{th}$ percentile error bars. All elevations in the high slope and high relief topographic transition zone (~1800–2500) have been filtered out.

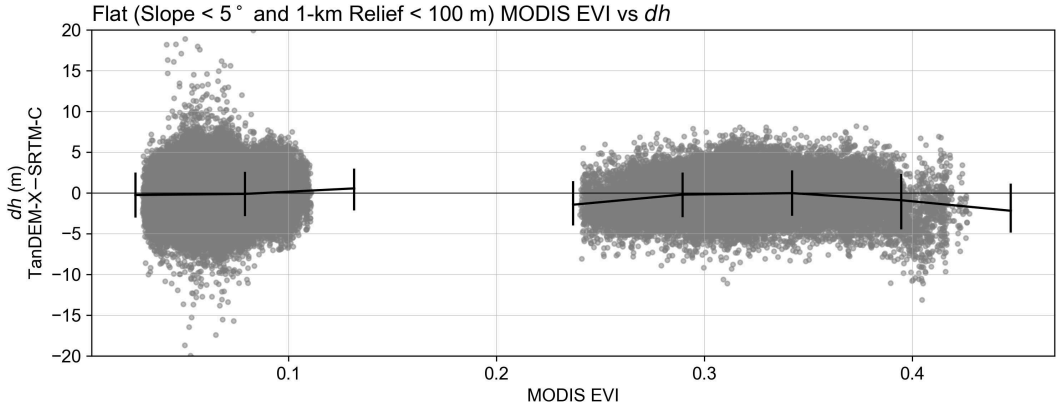


**Figure A9** – Vegetation versus *dh*. We use enhanced vegetation index (EVI) from MODIS product MOD13C1 (Huete et al., 1994) averaged over 14 years to track vegetation (Fig. 3.1C). Gray points are raw data and black line is median values with $10^{th}$ and $90^{th}$ percentile error bars. All elevations in the high slope and high relief topographic transition zone (~1800–2500) have been filtered out.

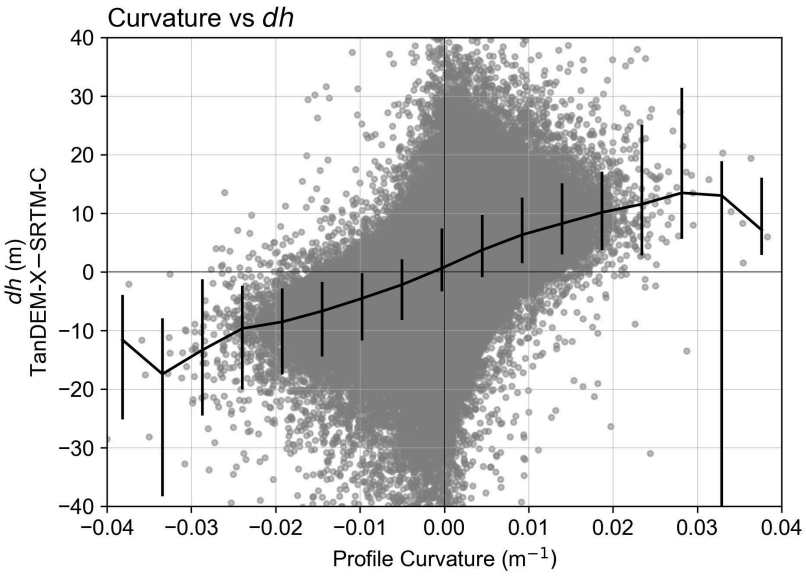


**Figure A10** – Terrain curvature versus *dh*. Gray points are raw data and black line is median values with $10^{th}$ and $90^{th}$ percentile error bars. Negative (positive) curvature values are valleys (ridges) corresponding to negative (positive) *dh* values where the SRTM-C is over- (under-) predicting heights.

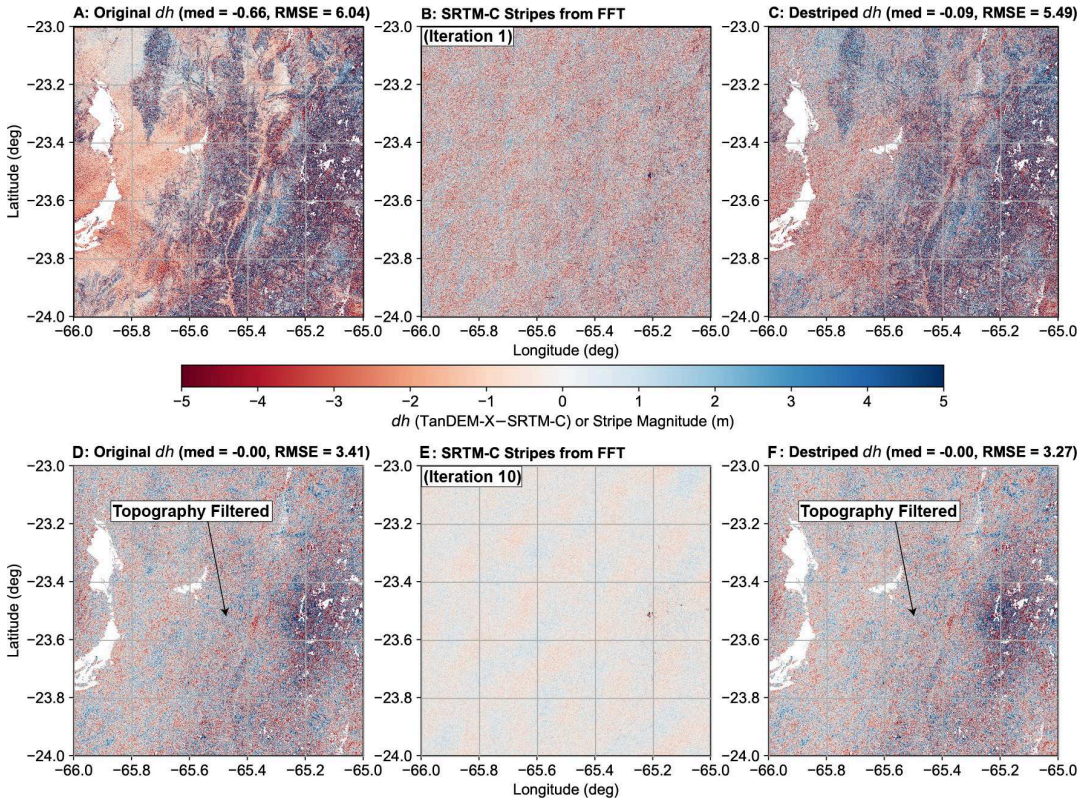


**Figure A11** – Two iterations of more aggressive FFT destriping from one tile (S 24°, W 66°), same as in Figure 3.3. Top panel (A–C) show first iteration and bottom panel (D–F) show tenth iteration. We note that while the striping and remaining patches caused by SRTM-C collection biases are removed using the more aggressive approach, the true topographic noise is also filtered, which is not the aim of our correction steps.

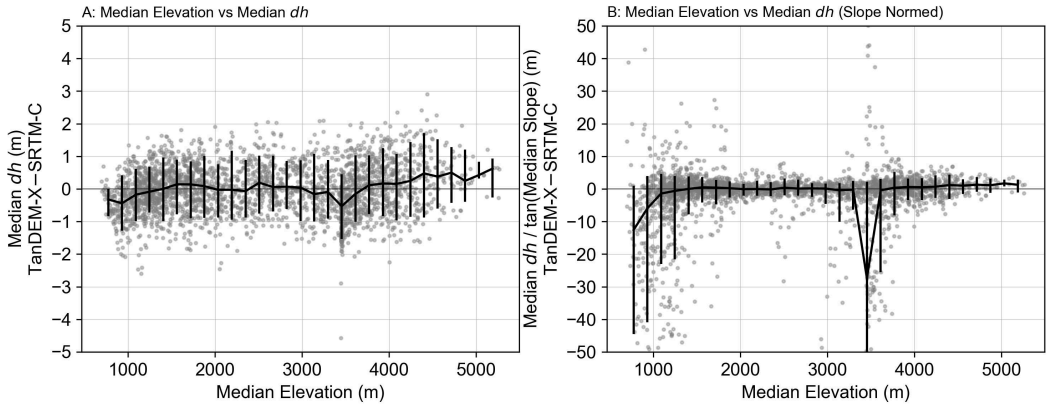


**Figure A12** – (A) Blocked median elevation versus median *dh*, and (B) same plot normalized by median slope in each block. Gray points are raw data and black line is median values with $10^{th}$ and $90^{th}$ percentile error bars. This is 2700 blocks of size 3.6×3.6 km over three tiles shown in Figure 3.4 (S 24–26°, W 66°). We do not note any distinct trends among the scatter.

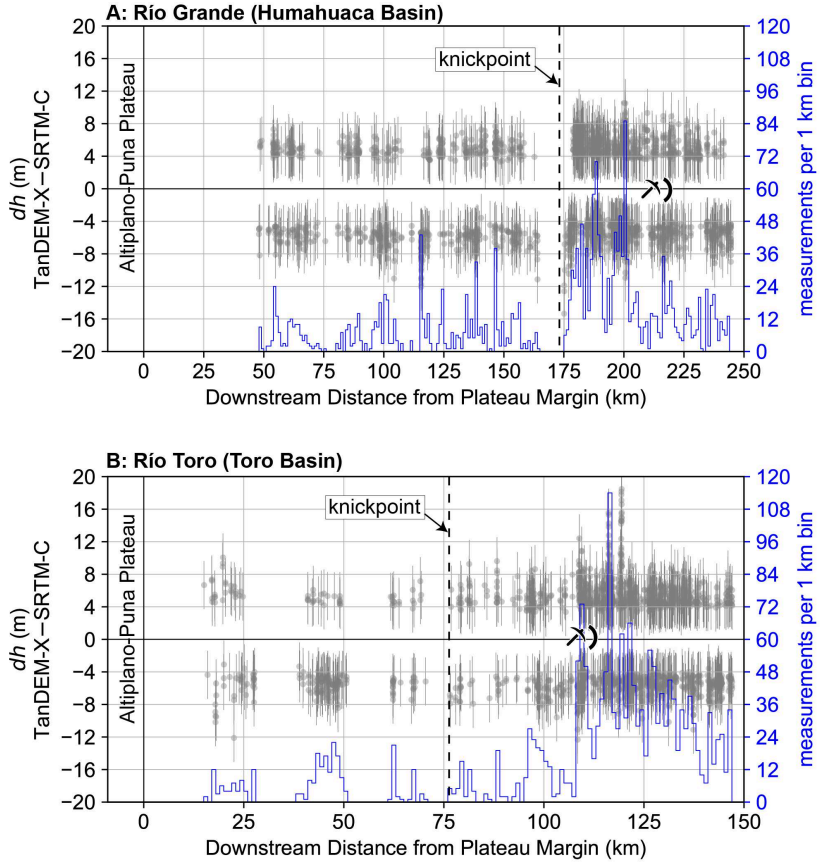


**Figure A13** – Point clouds (gray points) of potential *dh* signal (pixels outside of the range of expected noise) plotted with downstream distance for (A) Río Grande and (B) Río Toro. The left axis and points is identical to Figure 3.7. Error bars are RMSE from low slope (< 5°) terrain outside of the channel area. Rather than coloring each *dh* point cloud by relative density as in Figure 3.7, here we bin the measurements by 1-km reaches downstream and demonstrate the number of pixels on the right (blue) axis as a histogram. The location of the knickpoint is again indicated, demonstrating the approximately five to ten fold increase in measurements in the downstream reaches of both channels.

Appendix B

# Supporting Information for Chapter 4

## Additional Data Dimensions from Point Clouds

The results presented in Chapter 4 are similar to other studies segmenting grains from 2D imagery (e.g., Detert and Weitbrecht, 2012). This ignores the potential to exploit the third height dimension of the data from irregularly spaced SfM-MVS (or lidar) point clouds and associated DEMs. Many authors have already begun to look at patch-scale variance or roughness (e.g., Rychkov et al., 2012; Brasington et al., 2012) from point clouds on gravel-bed rivers to determine bulk characteristics, but this stops short of object detection and segmentation. Here, we briefly describe some of our own efforts to incorporate this additional information into *PebbleCounts*.

Our simplest approach was including the gridded DEM information, resampled to the same resolution as the orthomosaic. We inverted the elevation raster and flood-filled from the lowest points (tallest grains) using watershed approaches, similar to lidar tree-detection algorithms (e.g., Chen et al., 2006; Alonzo et al., 2015). For large, prominent grains with semi-spherical shapes, the flooded area was found to linearly increase until reaching the grain boundary, at which point the rate of area change jumped. We explored this break point as a potential segmentation tool for larger grains, but found that in the complex natural setting the shape of most grains is far from spherical, and furthermore, overlapping grains led to inconsistent behavior in the area breaks.

In an additional approach, we calculated both roughness and curvature at a variety of scales (5, 10, 50, 100 mm) directly from the point cloud using the open-source *CloudCompare* software (CloudCompare, 2018). This information was then gridded into a raster of the same resolution of the orthomosaic. While roughness could at times identify the smoother sand patches, it was difficult to discern between a sand patch and flat rock, and a color threshold on the orthoimagery was more successful. Curvature showed some spikes at grain boundaries, with the potential to aid in edge detection, however, we found that curvature was also high on intra-granular features.

In general, this analysis was complicated by vertical noise (scattering around a mean value) inherent to the SfM-MVS technique in the generation of dense point cloud data. In the field, for ~9 near-nadir photos taken from a height of ~4.5 m, the vertical standard deviation of points on a detrended flat surface (one of our coded targets) was found to be 1.7 mm for 13,014 points. On the other hand, in the perfect lab setting with 16 nadir+oblique photos from ~1.5 m, the detrended flat carpet around the pebbles achieved a standard deviation of 0.2 mm (33,371 points), similar to other SfM-MVS studies using large numbers of carefully collected images (e.g., Cullen et al., 2018; Verma and Bourke, 2019). These standard deviations from detrended flat surfaces represent a best-case scenario, whereas, in our field setting, the vertical

uncertainty on the complex, overlapping pebbles is likely higher. Such vertical noise is absent from the orthomosaics and limits the applicability of point clouds at these scales.

Ultimately, as the point cloud actually has a lower resolution (since it is based only on matched points) and more vertical noise than the orthomosaic (which exploits the full camera resolution), the imagery alone provided more detail. This is particularly important around grain edges needed for segmentation, which are not captured in nadir imagery alone, as shown in Figure B1. The lab setting resulted in point clouds with sufficient density and precision to identify individual grains with point-cloud processing tools. Thus, achieving higher quality SfM-MVS point clouds is possible, but only through more intense data collection during fieldwork.

Alternatively, lidar point clouds with distance measurements based on phase shifts have a lower standard deviation of ~1 mm in multiple settings and distances (up to ~300 m) and could allow more precise delineation using roughness and curvature calculations directly on the point cloud, however, such devices remain costly. Additionally, the development of affordable hyperspectral cameras with additional wavelengths will help in image segmentation in the spectral domain. To conclude, the potential for additional data dimension integration into pebble counting may be possible using higher dimensional object detection schemes, but, for the time-being, the orthoimagery alone provides satisfying results.

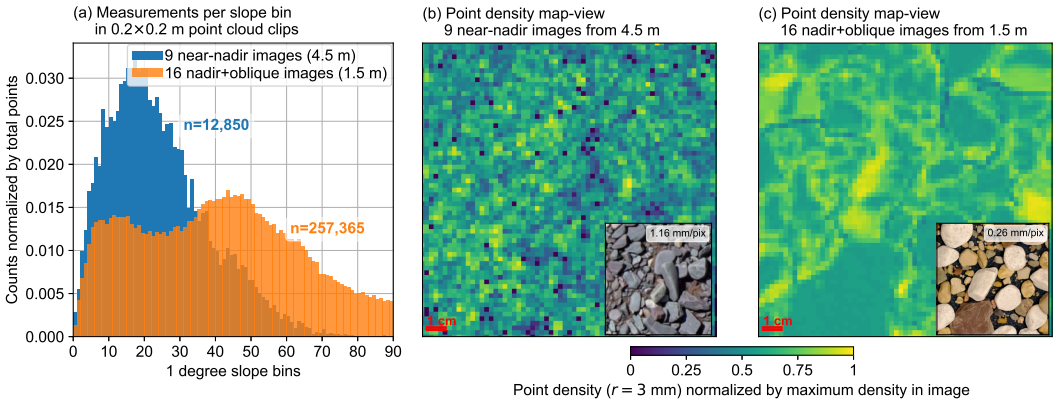


**Figure B1** – (a) Slope distribution in field (near-nadir) and experimental (nadir+oblique) point cloud clips. The point cloud slope was calculated in *CloudCompare* (CloudCompare, 2018) by first calculating the normals at each point using the 6 nearest neighbors and then extracting the dip of each normal. (b) Map-view of point density normalized by the maximum for the 9 near-nadir field images and (c) the same for the 16 nadir+oblique experimental images. Point density was calculated as the number of points in a radius of 3 mm. The clips were from a $0.2 \times 0.2$ m area, visually selected to have similar grain sizes and numbers of grains, shown in the inset images in (b) and (c). The average point density for the 16 nadir+oblique photo setting was 59 points/cm$^2$, whereas, in the field using 9 near-nadir photos the density was 17 points/cm$^2$. Note the higher point density on grain edges in (c) compared to (b), which are important for segmenting grains directly on the point cloud.

# Command-line Variables and Example Screenshots for *PebbleCounts*

Table B1 shows the command-line variables for *PebbleCounts* (KMS approach) and Table B2 shows the command-line variables for *PebbleCountsAuto* (AIF approach). Examples of the command-line interface and manual clicking steps are shown in Figure B2 and Figure B3, respectively.

**Table B1** – Command-line variable flags in *PebbleCounts* and their meaning. The default values are effective for most images.

| Variable Flag | Meaning (units) | Default Value(s) and Suggested Range |
|---|---|---|
| *im* | Image to run, including path to folder | No default |
| *ortho* | Georeferenced orthoimagery flag | No default, 'y' for orthoimagery, 'n' for nadir |
| *input_resolution* | Input resolution if not orthoimage (mm) | No default, calculate from eq. 4.1 |
| *subset* | Interactively subset image | Default no ('n') |
| *sand_mask** | Name, including path, to a sand mask if one already exists | No default |
| *otsu_threshold** | Percentage of Otsu value to threshold shadows by (percentage of 100) | No default, suggested value of 50 |
| *maxGS** | Expected maximum a-axis grain size (m) | Default 0.3 |
| *cutoff** | Minimum b-axis length to be counted (pixels) | Default 20, can be raised |
| *min_sz_factors** | Factors to multiply *cutoff* at each scale, used to cleanup masks for easier clicking | Default [50, 5, 1] for three scales (large to small) for ∼1 mm/pixel imagery, double for < 0.8 mm/pixel |
| *win_sz_factors** | Factors to multiply *maxGS* by at each scale | Default [10, 3, 2] for three scales (large to small), can be changed ±0.5–1.5 to get more or less windows |
| *improvement_ths** | Improvement threshold values that tell k-means when to halt (fraction of 1) | Default [0.01, 0.1, 0.1] for three scales (large to small), can be varied from 0.01–0.2 |
| *coordinate_scales** | Fraction to scale *x,y* coordinates (fraction of 1) | Default [0.5, 0.5, 0.5] for three scales (large to small), can be varied from 0.3–0.7 |
| *overlaps** | Fraction of overlap between windows (fraction of 1) | Default [0.5, 0.3, 0.1] for three scales (large to small), can be varied from 0–0.5 at each scale |
| *first_nl_denoise** | Strength of first non-local means denoising | Default 5, can be varied ±1 |
| *nl_means_chroma_filts** | Strength of windowed non-local means denoising | Default [3, 2, 1] for three scales (large to small), can be varied ±1 |
| *bilat_filt_szs** | Size of bilateral filtering windows (pixels) | Default [9, 5, 3] for three scales (large to small), can be varied from 3–9 |
| *tophat_th** | Upper percentile threshold to take from top-hat filter for edge detection (fraction of 1) | Default 0.9, can be varied from 0.8–0.95 |
| *sobel_th** | Upper percentile threshold to take from sobel filter for edge detection (fraction of 1) | Default 0.9, can be varied from 0.8–0.95 |
| *canny_sig** | Canny filtering sigma value for edge detection | Default 2, can be varied from 1–2 |
| *resize* | Value to resize windows by (fraction of 1) | Default 0.8, can be varied from 0.5–0.99 if you want a smaller (0.5) or larger (0.99) pop-up window |

*Influence on results.

**Table B2** – Command-line variable flags in *PebbleCountsAuto* and their meaning. The default values are effective for most images.

| Variable Flag | Meaning (units) | Default Value(s) and Suggested Range |
|---|---|---|
| *im* | Image to run, including path to folder | No default |
| *ortho* | Georeferenced orthoimagery flag | No default, 'y' for orthoimagery, 'n' for nadir |
| *input_resolution* | Input resolution if not orthoimage (mm) | No default, calculate from eq. 4.1 |
| *subset* | Interactively subset image | Default no ('n') |
| *sand_mask** | Name, including path, to a sand mask if one already exists | No default |
| *otsu_threshold** | Percentage of Otsu value to threshold shadows by (percentage of 100) | No default, suggested value of 50 |
| *cutoff** | Minimum b-axis length to be counted (pixels) | Default 20, can be raised |
| *percent_overlap** | Maximum allowable overalp between neighboring ellipses for filtering suspect grains (percentage of 100) | Default 15, can be varied from 5–30 |
| *misfit_threshold** | Maximum allowable misfit between ellipse and grain mask for filtering suspect grains (percentage of 100) | Default 30, can be varied from 10–50 |
| *min_size_threshold** | Minimum area of grain, used to clean the mask (pixels) | Default 10 for ~1 mm/pixel imagery, 40 for < 0.8 mm/pixel |
| *first_nl_denoise** | Strength of first non-local means denoising | Default 5, can be varied $\pm 1$ |
| *tophat_th** | Upper percentile threshold to take from top-hat filter for edge detection (fraction of 1) | Default 0.9, can be varied from 0.8–0.95 |
| *sobel_th** | Upper percentile threshold to take from sobel filter for edge detection (fraction of 1) | Default 0.9, can be varied from 0.8–0.95 |
| *canny_sig** | Canny filtering sigma value for edge detection | Default 2, can be varied from 1–2 |
| *resize* | Value to resize windows by (fraction of 1) | Default 0.8, can be varied from 0.5–0.99 if you want a smaller (0.5) or larger (0.99) pop-up window |

*Influence on results.


**Figure B2** – Example of command-line and pop-up interface for *PebbleCounts*. (a) Interactive Otsu thresholding using percentage of Otsu value and yes ('y') or no ('n') confirmation. (b) Interactive color masking by yes ('y') or no ('n') and resulting color mask after selection. (c) K-means clustering and pop-up window for pebble selection by left clicking, with black arrows measured in final output and red arrows ignored after right-click removal (see Fig. B3).

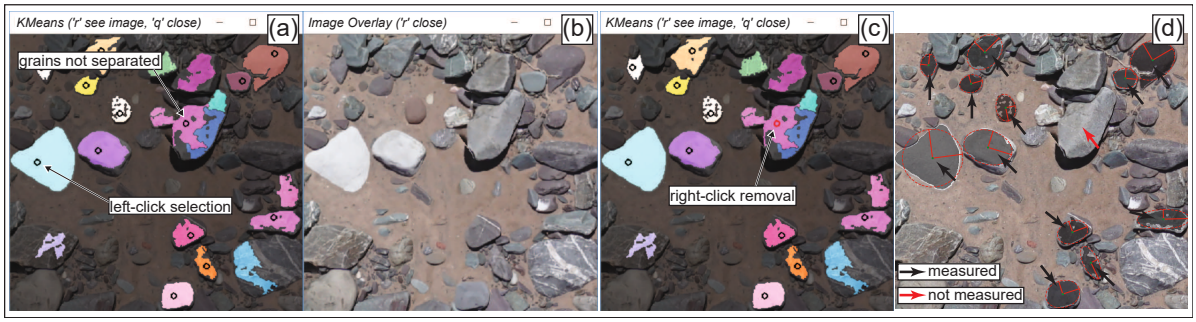


**Figure B3** – Clicking tutorial continued from Figure B2c. Following k-means clustering at each scale a mask overlaid on the original image is presented (a), and grains are selected by a left click anywhere in the segmented area, resulting in a black circle at the click location. When clicking is finished the mask is closed by pressing 'q'. To view the original unmasked image the user may press 'r' (b). Using this switching the user can see which grains are poorly delineated and remove the last click with a right click on the mouse (c). The original black circle selection turns to red to signify this grain is off and will not be measured in the final output (d).

# Resampling and Parameter Selection in AIF Approach

Figure B4 demonstrates the percentage of grains with a match found in the AIF approach when increasing resampling from a factor of 0.6–2.6 by 0.1 steps using Lanczos resampling (Lanczos, 1950). As the resampling factor increases, there is progressive reduction in the number of found grains after filtering, therefore we selected the original resolution (resampling factor of one). Figure B5 and Figure B6 demonstrate two cases where the resampling slightly improved the resulting grain-size distribution. Both images were of relatively low quality with significant blurring and the presence of many weak edges between grains of similar color.

We selected a maximum percent misfit between the ellipse and grain of 30% as the $90^{th}$ percentile of misfits for the KMS approach was 30%. Furthermore, we allowed a maximum overlap between neighboring ellipses of 15%, visually selected to minimize overlapping grain measurement and over-segmentation of discrete grains. For the higher resolution imagery it was necessary to use a lower sobel and top-hat threshold (0.85), since we consider all the edges at once in the AIF approach, rather than in a windowed subset as in the KMS approach, and many edges are not found when using the 0.9 threshold given the increased number of pixels under consideration.

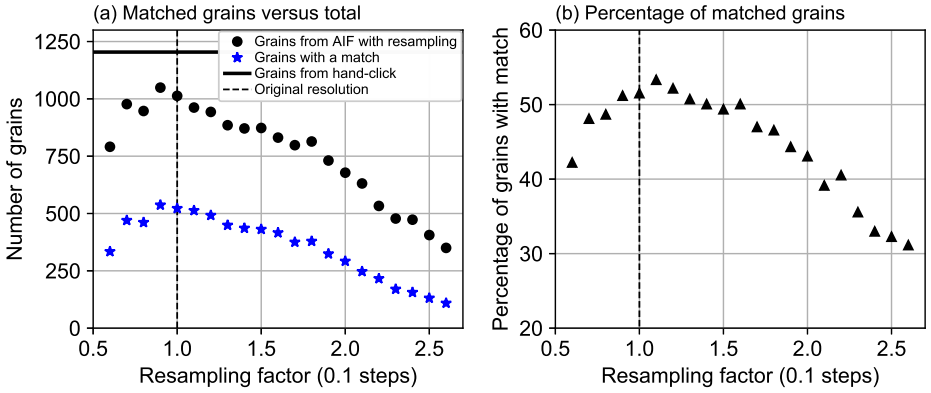


**Figure B4** – Matching grains found in each filtered mask versus the resampling factor (where 1 is the original image) for the ~1.16 mm/pixel resolution images. Matches are defined as an AIF grain within 5 pixels of the hand-clicked line or the KMS grain centroid and with a 1 cm maximum b-axis difference between the AIF grain and the match.

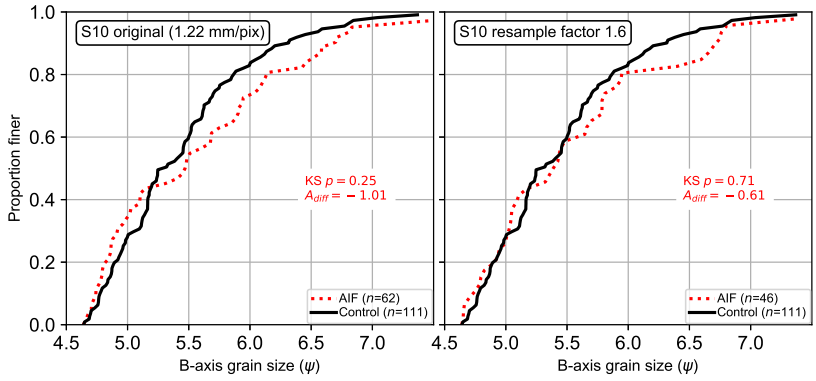


**Figure B5** – Slight improvement (increase in $p$ and decrease in $A_{diff}$) in result using a 1.6-times resampling factor prior to running the AIF algorithm for the difficult (somewhat blurry, weak edges) S10 orthoimage.

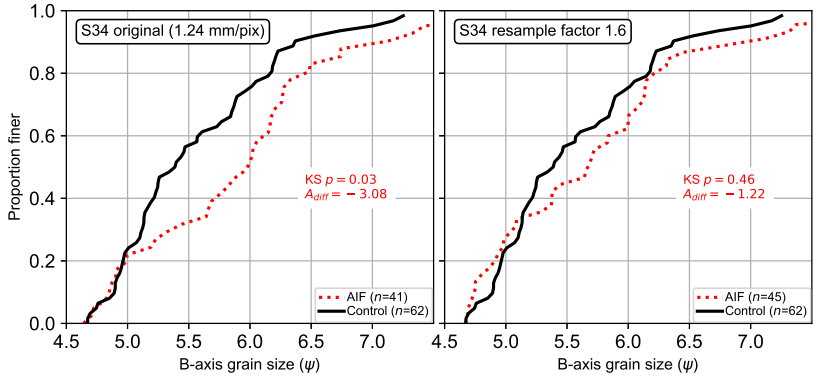


**Figure B6** – Slight improvement (increase in $p$ and decrease in $A_{diff}$) in result using a 1.6-times resampling factor prior to running the AIF algorithm for the difficult (very blurry, weak edges) S34 orthoimage.

# Agisoft Orthomosaic Generation

*Agisoft* (Agisoft, 2019) processing was carried out in the following steps for the indoor handheld imagery (with field-gathered mast imagery differences in parentheses following the step):

1. Image quality detection and the exclusion of photos with quality metric < 0.7. This step analyzes pixel contrast to estimate sharpness with values ranging from 0/blurred to 1/sharp. We found 0.7 to be a sufficient lower cutoff upon visual inspection of results.

2. Detection of 12-bit coded targets in the remaining photos, with two targets placed at each of the four corners of the area and ensuring that the diameter of the printed targets' center circle was limited to 10–30 pixels in image resolution for successful automated detection.

3. Input of scale for the orthomosaic output, provided by the distances between the targets at each corner, resulting in four distance measurements, with 0.5 mm accuracy using a ruler with cm and mm demarcations. (For the field images: The scale was provided by the XYZ coded target locations in UTM zone 19S, WGS84 ellipsoidal datum.)

4. Photo alignment at high quality with a 40,000 key-point and 2000 tie-point limit.

5. Dense cloud generation from the aligned photos at the medium output and with moderate depth filtering. Given the high quality of the photos more aggressive options did not improve results. (For the field images: Given the increased complexity of the setting and imperfect photo collection, the dense point cloud was generated at high quality with aggressive depth filtering.)

6. DEM building from the dense cloud with default settings in a local coordinate system. (For the field images: The DEMs and orthomosaics were also output in UTM zone 19S projections, providing undistorted pixels with resolution in m/pixel.)

7. Generation of an orthomosaic using the DEM for orthorectification at the default settings.

8. Output of the orthomosaic to a GeoTiff file with resolution provided in m/pixel.

# KMS and AIF Results Separated by Site

Here we show all of the results (following 20-pixel truncation) for each of the 12 sites in Figure B7. These results are aggregated in curves shown in Chapter 4 Figure 4.11 and a comparison of the individual percentiles of interest is shown in Chapter 4 Figure 4.12.

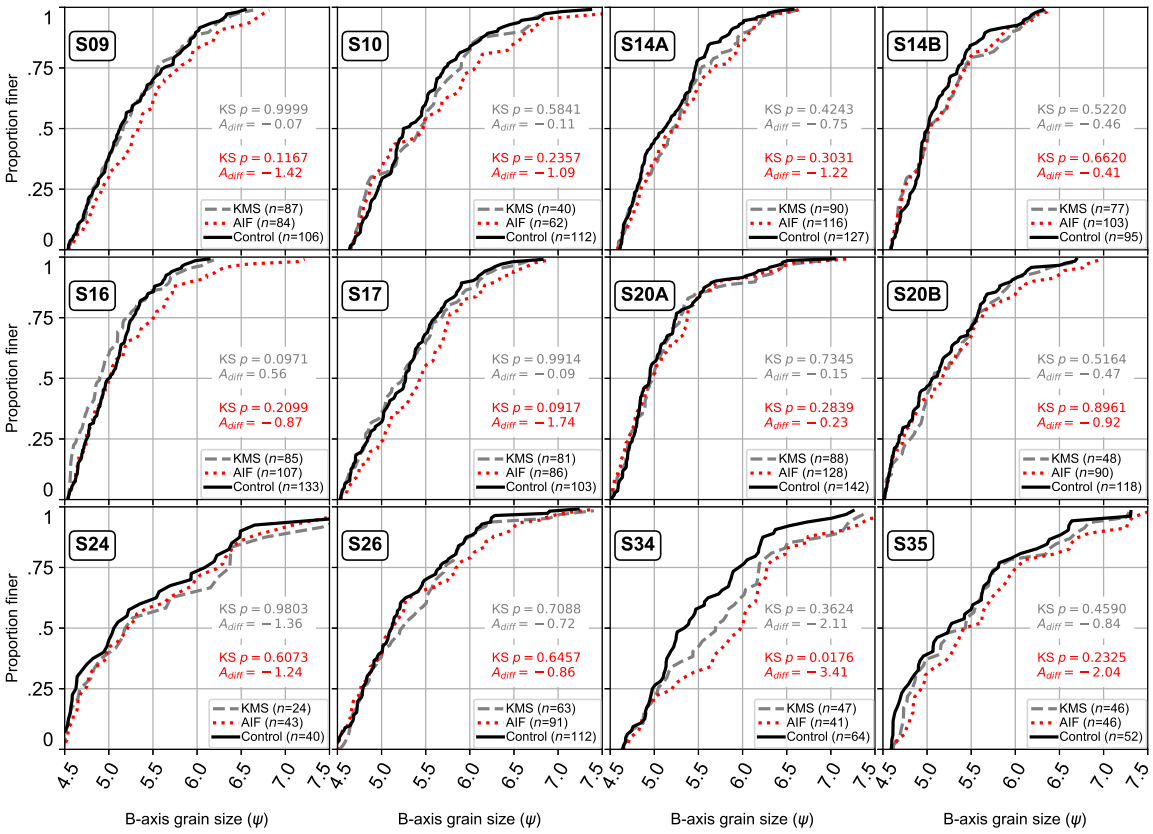


**Figure B7** – Comparison of 20-pixel truncated grain-size distributions between hand-clicked control (black line), KMS *PebbleCounts* (gray, dashed line), and AIF *PebbleCountsAuto* (red, dotted line) for the 12 × ~1.16 mm/pixel control sites. In corresponding colors are the $p$-value results of a KS-test and the $A_{diff}$ approximate integral between the curves for each approach versus the control data. The legend indicates the number of grains ($n$) making up each curve. See Figure 4.6b in Chapter 4 for sites.

# Misidentification in the AIF Approach

Figure B8 demonstrates remaining issues with the AIF approach in a few map-view examples. On a grain-by-grain basis, there are many inaccuracies falling into three main categories: over-segmentation of grains with internal edges and the selection of each segment as a separate grain, under-segmentation and merging of neighboring grains that have weak edges sometimes caused by image blur, and misidentification of non-grain objects or clusters of small grains. It is clear from this analysis that caution must be used when interpreting AIF results, particularly in complex or blurry images.

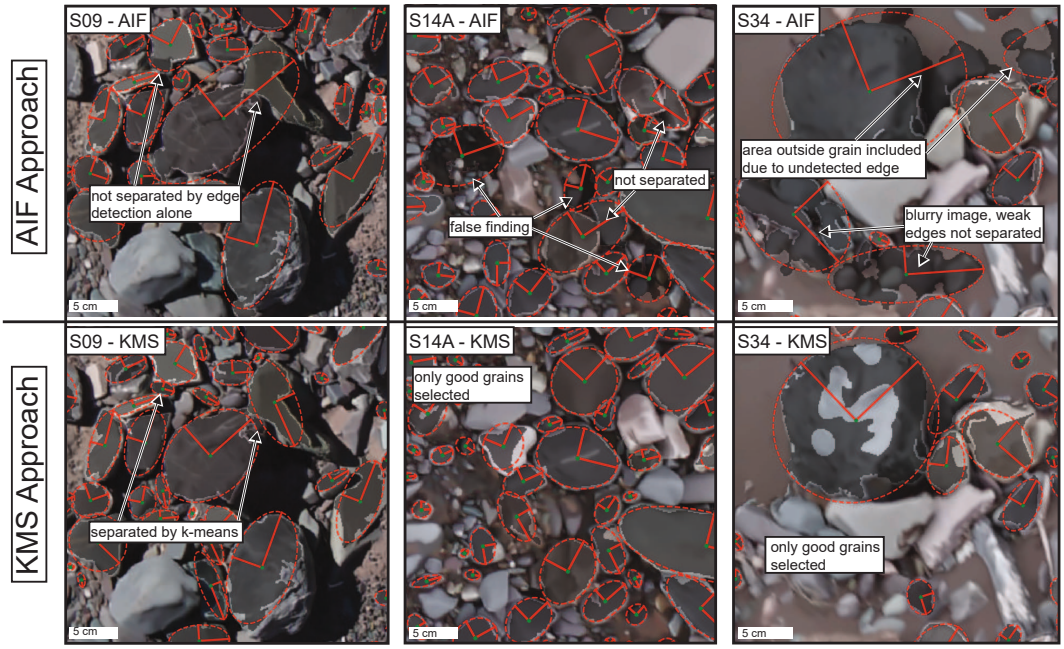


**Figure B8** – Resulting delineated grains using the AIF *PebbleCountsAuto* function (top row) versus the same area from the KMS *PebbleCounts* function (bottom row). Labels indicate the issues with the AIF results and improvement in KMS results. Note the poor results for the blurry image on the right (S34).

Appendix C

# Supporting Information for Chapter 5

## Contents of Data Repository

This appendix contains a table (Table C1) with the SAR datasets used, followed by a number of figures referenced in Chapter 5 (Figures C1–C14).

**Table C1** – SAR datasets.

| **Sensor (short name)** | **Wavelength** | **Polarization** | **Dates (YYYY-MM-DD) (ASC=Ascending DSC=Descending)** | **Mean ± standard deviation of local incidence angle over channel** |
|---|---|---|---|---|
| TerraSAR-X/ TanDEM-X (TSX/TDX) | X-band (3.1 cm) | HH | 2011-08-02 (DSC)<br>2011-09-15 (DSC)<br>2012-03-20 (DSC)<br>2013-02-24 (DSC)<br>2013-03-07 (DSC)<br>2013-10-03 (ASC)<br>2014-03-05 (DSC)<br>2014-01-10 (ASC)<br>2015-04-28 (ASC)<br>2015-07-13 (DSC)<br>2015-09-07 (ASC)<br>2015-09-28 (DSC) | $36 \pm 16$ |
| Sentinel-1 (S1) | C-band (5.6 cm) | VV, VH | 2017-06-10 (DSC)<br>2017-07-16 (DSC)<br>2017-11-17 (ASC)<br>2017-12-07 (DSC)<br>2017-12-11 (ASC)<br>2018-01-16 (ASC)<br>2018-02-21 (ASC)<br>2018-03-11 (ASC)<br>2018-05-16 (ASC)<br>2018-06-15 (ASC)<br>2018-07-11 (DSC)<br>2018-07-15 (ASC)<br>2018-09-01 (ASC)<br>2018-11-12 (DSC)<br>2019-03-12 (ASC) | $39 \pm 14$ |
| ALOS-2 PALSAR-2 (ALOS2) | L-band (22.9 cm) | HH, HV | 2015-09-23 (ASC)<br>2016-02-10 (ASC)<br>2016-09-21 (ASC) | $31 \pm 7$ |

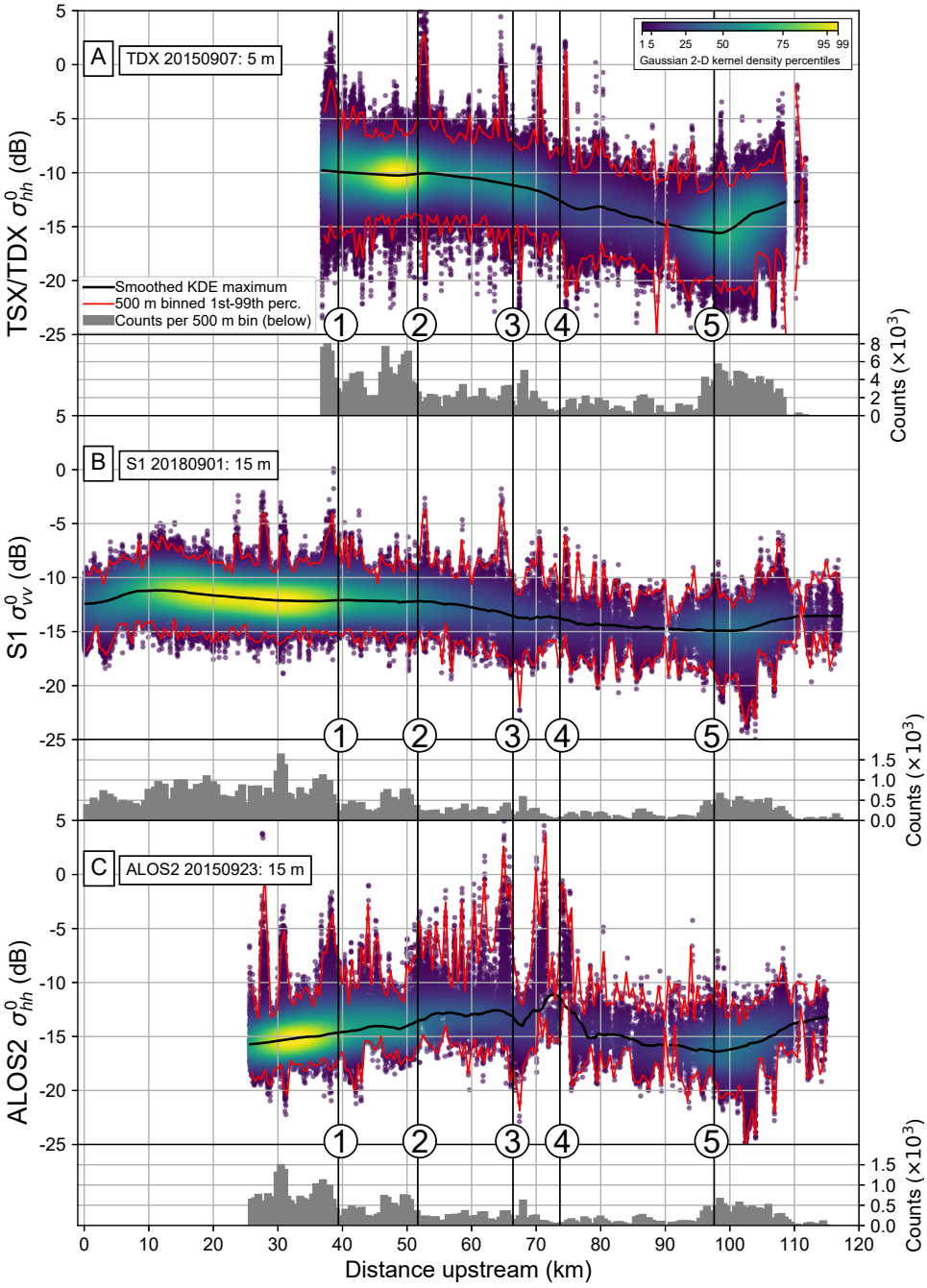


**Figure C1** – Same as Figure 5.6C in Chapter 5, now showing two additional scenes from the other sensors. Note greater number of counts per bin for the TSX/TDX data in (A) since the resolution is 5 m rather than 15 m. The tectono-geomorphic transitions (1–5) correspond to the map-view image in Chapter 5 Figure 5.1B.

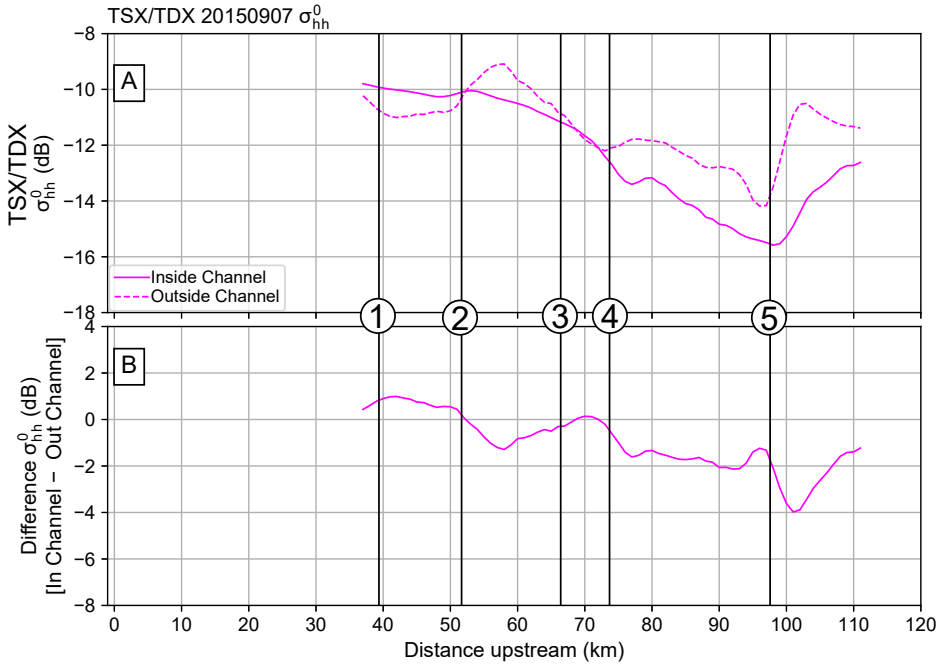


**Figure C2** – KDE trendline differencing for one TSX/TDX $\sigma^0_{hh}$ scene (20150907). (A) shows the inside-channel and outside-channel trendlines extracted from the full NDVI-masked (0.15 inside-channel and 0.35 outside-channel threshold) raw $\sigma^0$ point clouds with the difference (inside-outside) shown in (B). The tectono-geomorphic transitions (1–5) correspond to the map-view image in Chapter 5 Figure 5.1B.

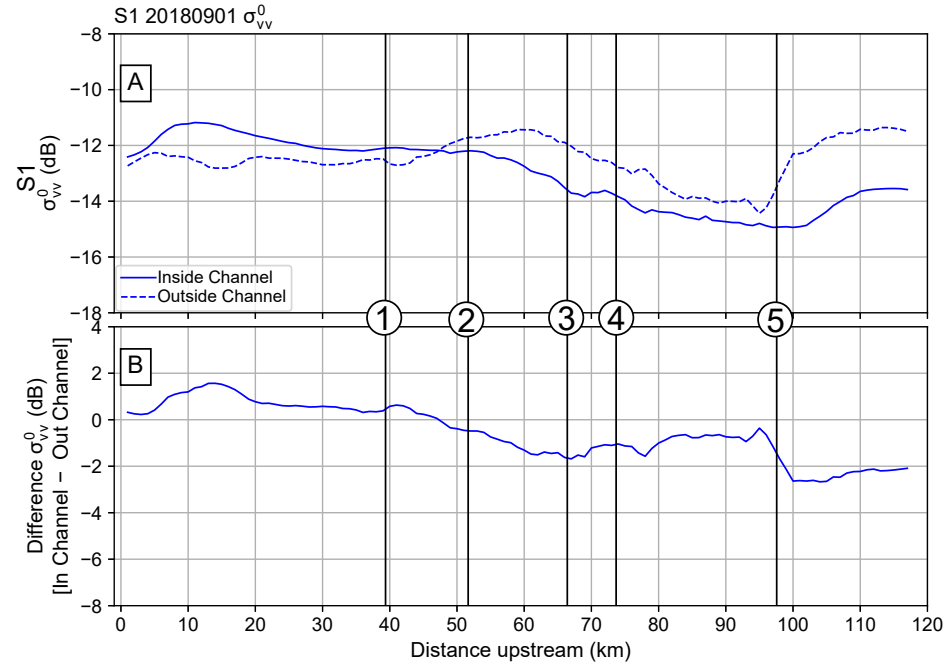


**Figure C3** – KDE trendline differencing for one S1 $\sigma^0_{vv}$ scene (20180901). See Figure C2 caption for description of (A) and (B).

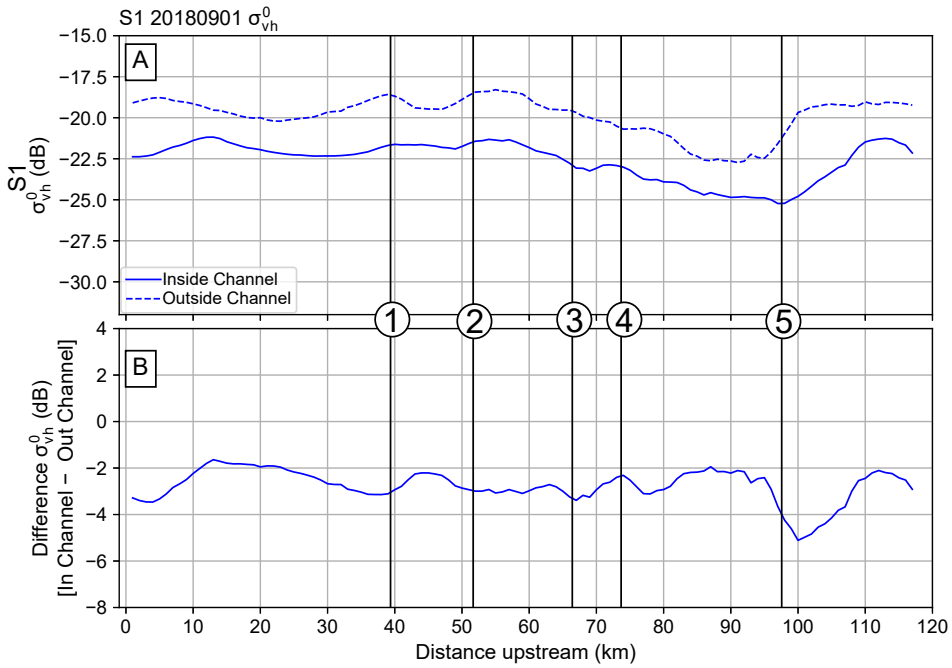


**Figure C4** – KDE trendline differencing for one S1 $\sigma^0_{vh}$ scene (20180901). See Figure C2 caption for description of (A) and (B).

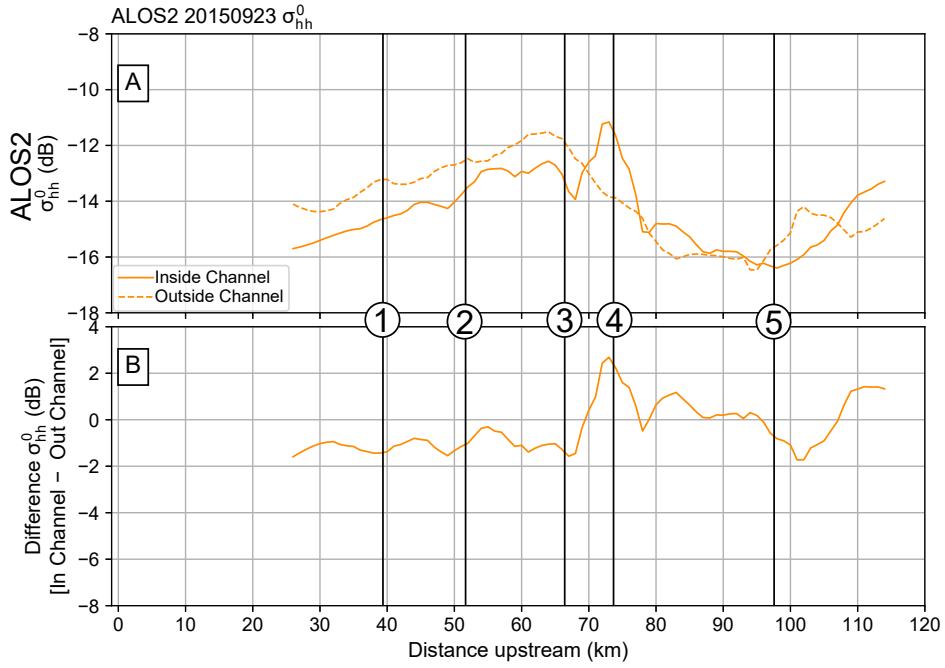


**Figure C5** – KDE trendline differencing for one ALOS2 $\sigma^0_{hh}$ scene (20150923). See Figure C2 caption for description of (A) and (B).

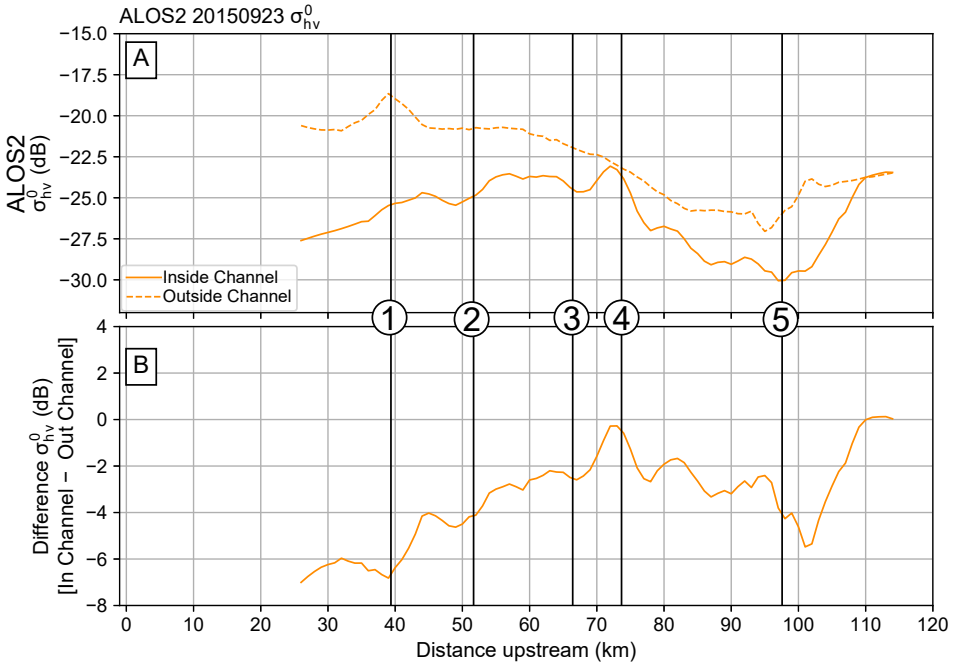


**Figure C6** – KDE trendline differencing for one ALOS2 $\sigma^0_{hv}$ scene (20150923). See Figure C2 caption for description of (A) and (B).

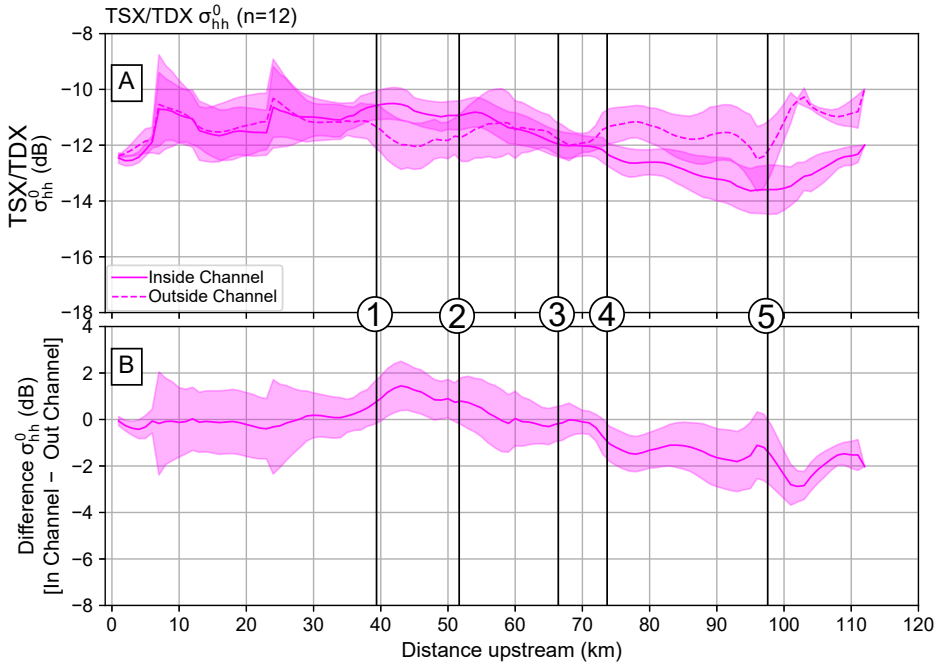


**Figure C7** – Integrated trendline for all TSX/TDX $\sigma^0_{hh}$ scenes. (A) shows the mean and standard deviation envelope for all the inside and outside-channel trendlines with the resulting differenced line in (B), where propagation of standard deviation from the two trendlines in (A) is used to get the envelope. The tectono-geomorphic transitions (1–5) correspond to the map-view image in Chapter 5 Figure 5.1B.

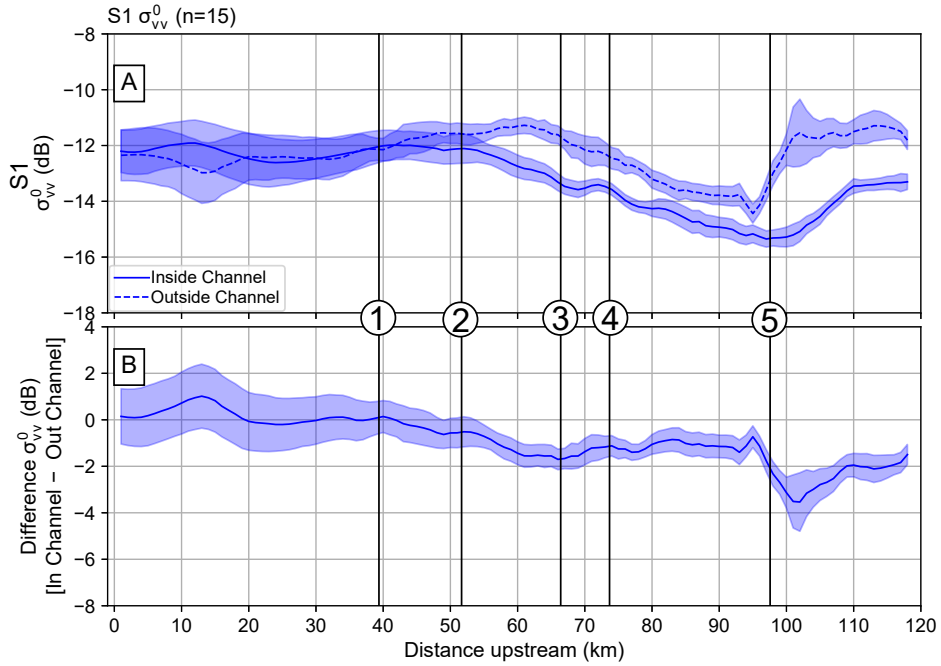


**Figure C8** – Integrated trendline for all S1 $\sigma^0_{vv}$ scenes. See Figure C7 caption for description of (A) and (B).

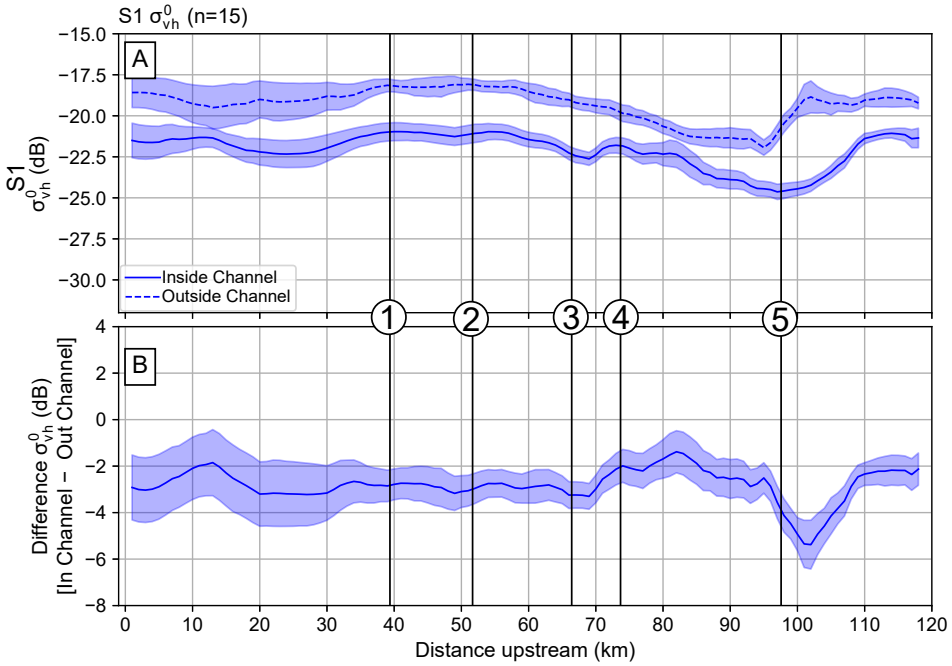


**Figure C9** – Integrated trendline for all S1 $\sigma^0_{vh}$ scenes. See Figure C7 caption for description of (A) and (B).

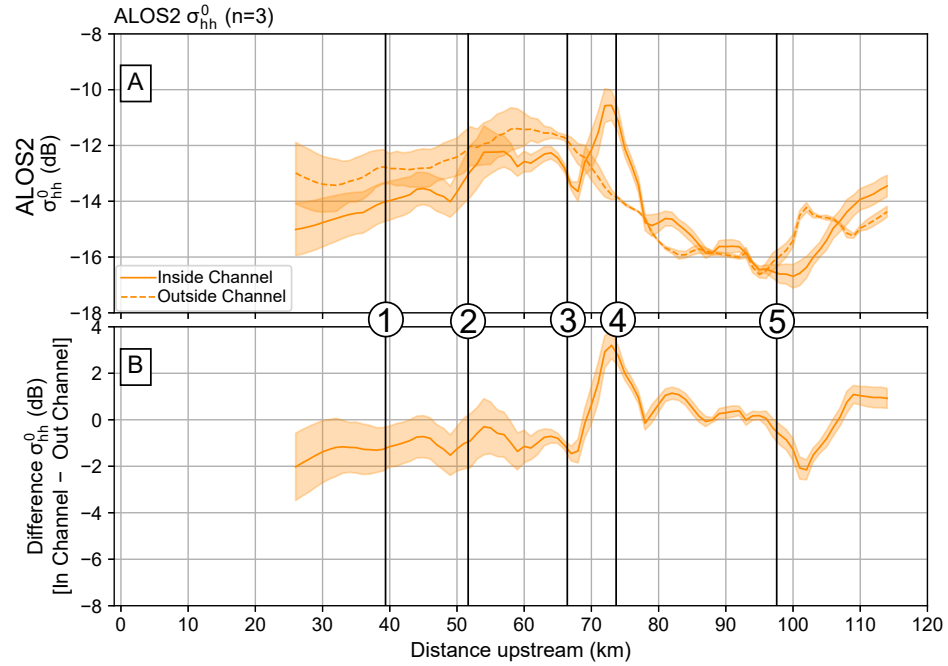


**Figure C10** – Integrated trendline for all ALOS2 $\sigma^0_{hh}$ scenes. See Figure C7 caption for description of (A) and (B).

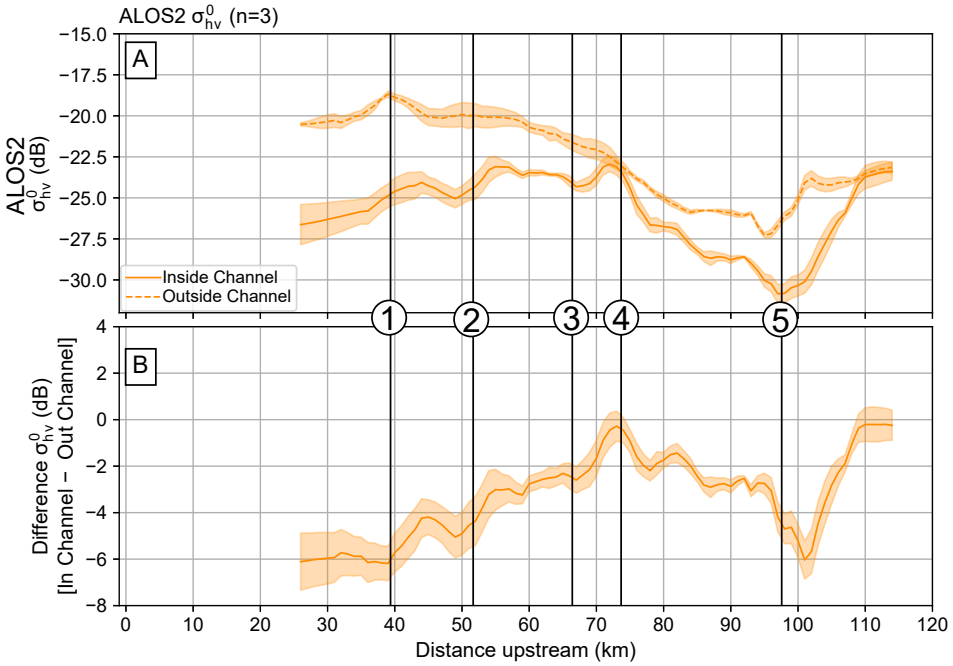


**Figure C11** – Integrated trendline for all S1 $\sigma^0_{hv}$ scenes. See Figure C7 caption for description of (A) and (B).

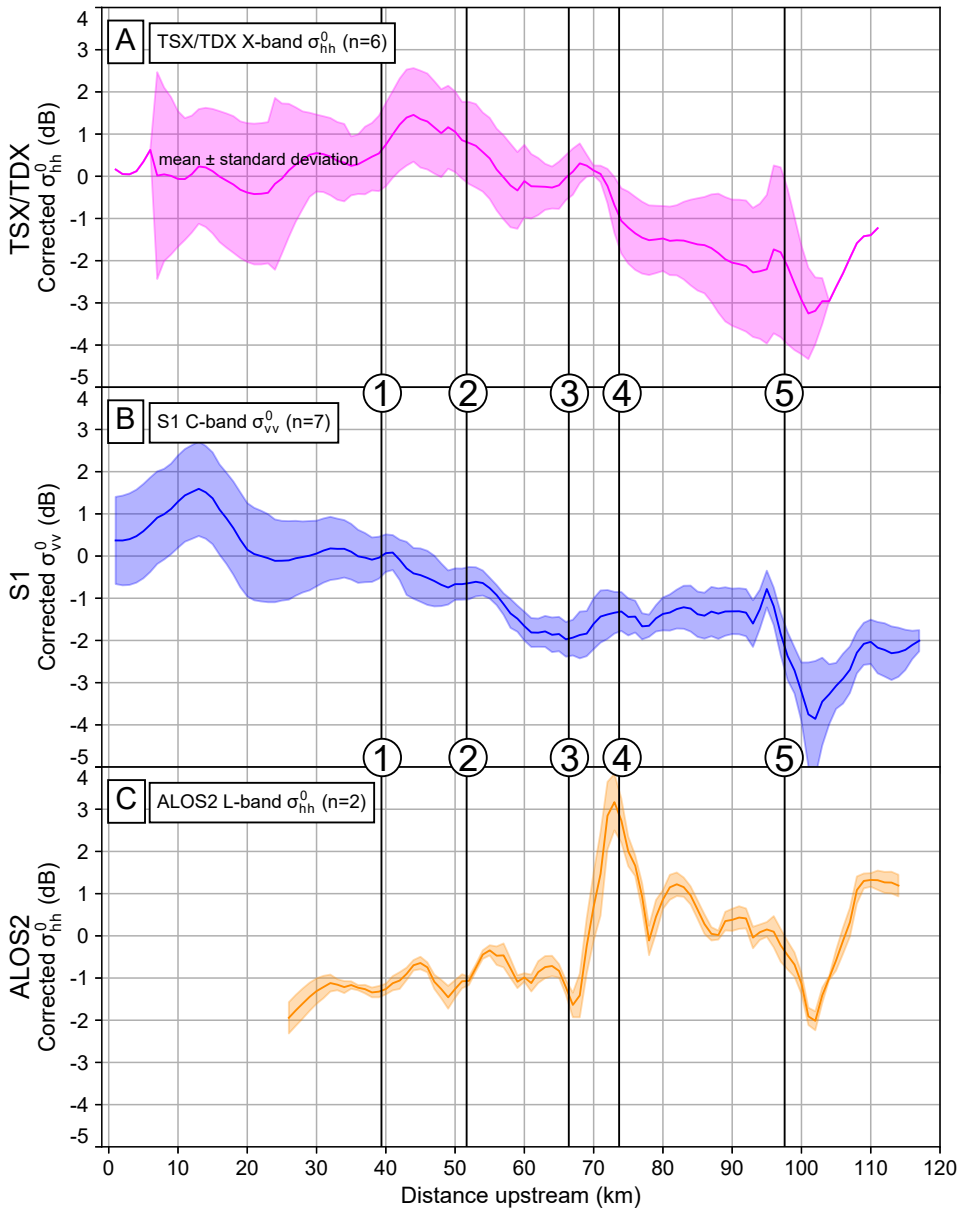


**Figure C12** – Same as Chapter 5 Figure 5.7, but here separating the scenes into only the dry season (April, May, June, July, August, and September). We note the same trend regardless of season.

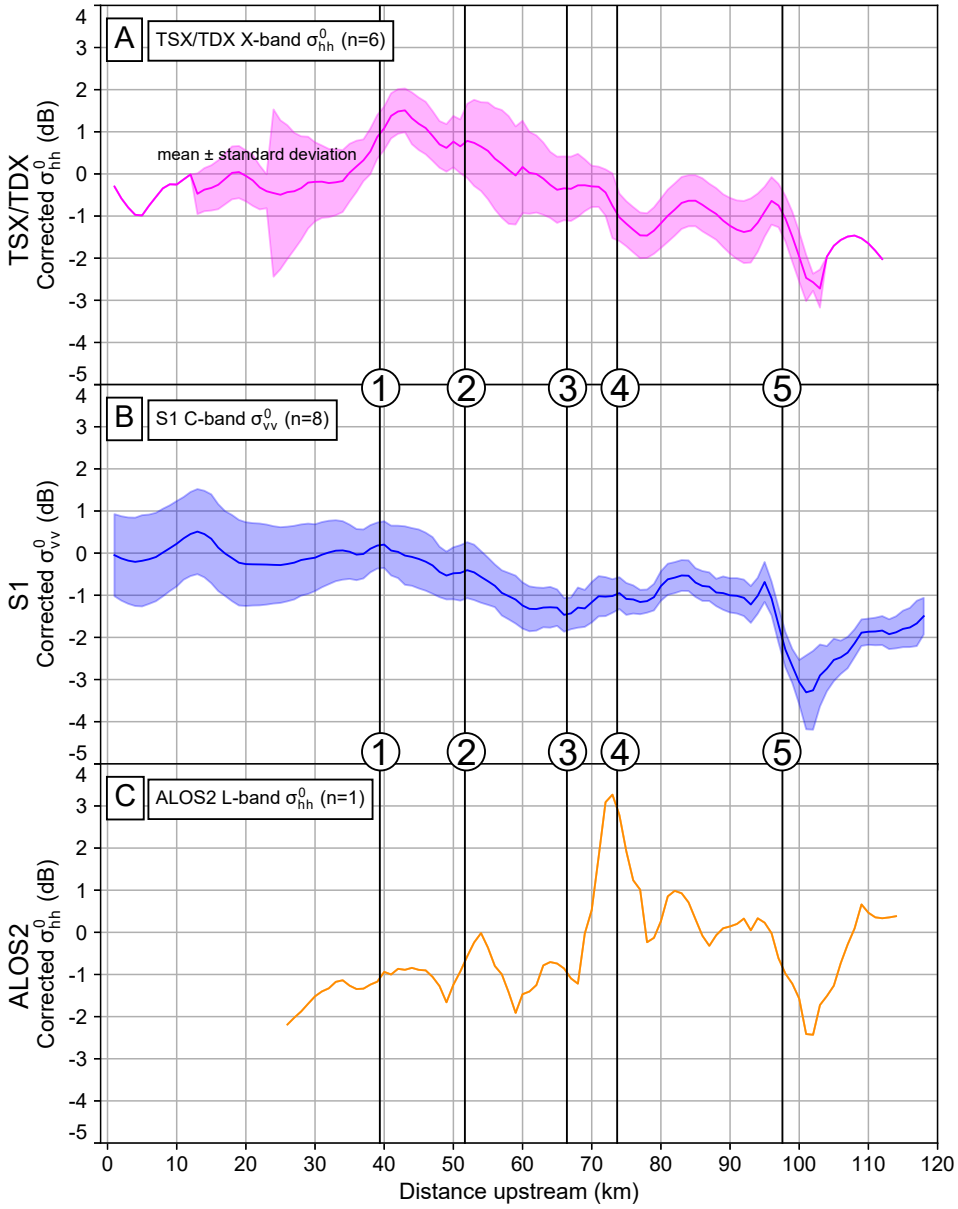


**Figure C13** – Same as Chapter 5 Figure 5.7, but here separating the scenes into only the wet season (October, November, December, January, February, and March). We note the same trend regardless of season.

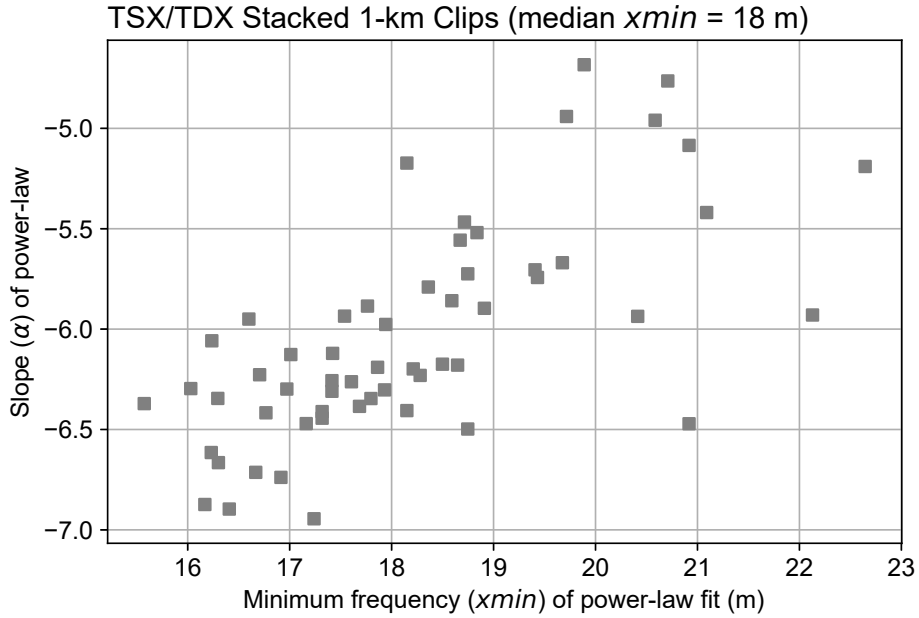


**Figure C14** – Slope of short-wavelength power-law fit on the 59 TSX/TDX stacked 1-km channel clips versus the minimum frequency ($xmin$) chosen as the best fit from the minimum KS-test statistic. Note the approximately linear trend of decreasing slope (shallowing) with increasing $xmin$. Despite this trend, the range of the $xmin$ values only covers ~7 m (< 2 TSX/TDX 5-m pixels). Given this tight range, we always fix the $xmin$ at the median value of these best-fit results (18 m) to compare power-law slopes regardless of minimum frequency of the fit.

# Declaration

This work has not been submitted to any other institution of higher education, and was prepared independently by Benjamin Purinton.

Potsdam, 23. Oct. 2019

Benjamin Purinton