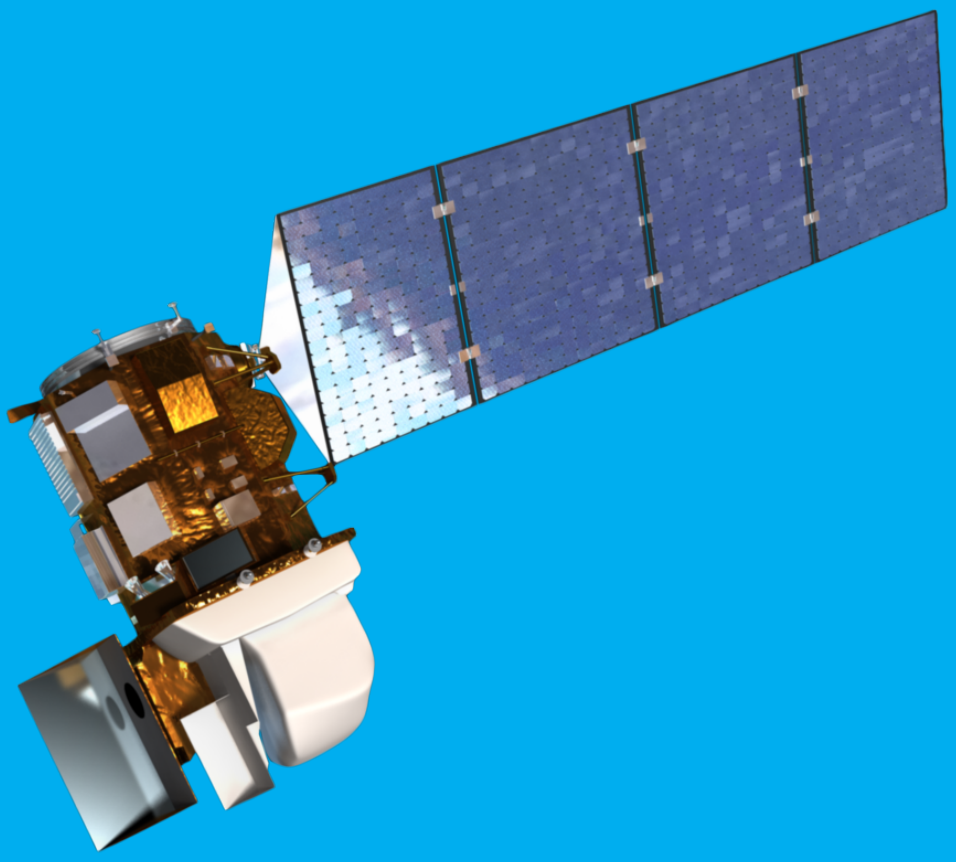
Design and implementation of a novel method for creating glacial velocity time series for the Himalayan region using the optical Landsat database.

S. Pinson



TU Delft

Design and implementation of a novel method for creating glacial velocity time series for the Himalayan region using the optical Landsat database.

S. Pinson

by

to obtain the degree of Master of Science
at the Delft University of Technology,
to be defended publicly on Tuesday December 19, 2017 at 14:00 AM.

Student number:	4105699	
Project duration:	December 1, 2016 – December 19, 2017	
Thesis committee:	Dr. R. Lindenbergh,	TU Delft, supervisor
	Prof. DR. M. Menenti,	TU Delft
	Prof. Dr. N. Gourmelen,	Edinburgh University
	Dr. S. LHermitte,	TU Delft

TU Delft

Preface

This thesis presents a step towards better glacial velocity time series. The research started of as an 4.5 month internship at the University of Edinburgh at the School of Geosciences, but evolved into a full thesis that took the better part of a whole year to complete. With this research my MsC Geoscience and Remote Sensing comes to an end. This relatively new master, the people I worked with in the master and the subjects discussed in this master at least inspired me to change my view on many subjects regarding the Earth system and sustainability, for which I am grateful. I hope that the fruit of this research can be used in future research and contributes towards a better understanding of glaciers around the world.

With that, I would like to thank all the people who contributed to this research. Most of all my supervisor from Edinburgh University, Noel Gourmelen, who offered me the chance to study in (the amazing city of) Edinburgh and work on this thesis. His technical knowledge of the subject was of great help, but more so I enjoyed the company of the research group in Edinburgh. The supervision back in the Netherlands from Roderik Lindenbergh was also of great help, for which I would like to thank him. His advise on report structure and writing and effort was of great help. Furthermore I would like to thank Prof Massimo Menenti for his clear advice on the report structure and Steph Lhermitte for taking the time to provide some very useful feedback, while in Antarctica. Finally, I would also like to thank Eric Tol, Amber van Hauwermeiren, Lars Keuris and my father Jan Pinson for reading and checking parts of my thesis and supporting my work.

Delft, December 2017
It's nothing

Abstract

Glaciers play an important role in sea level predictions and are an important supplier of fresh water. Understanding the physics and dynamics of glaciers is important in local and global climate predictions and predictions of fresh water supplies. The Himalayas are the biggest storage of fresh water outside the polar regions with many different local climates and glaciers. One important glacier parameter used in dynamic studies is the flow velocity, as the flow velocity is used as a boundary condition in mass balance and run-off models. The flow velocity of a glacier is relatively easy to measure on a large scale using satellite missions and the increased coverage of satellite missions provides the means to large scale velocity monitoring. Robust methods for measuring large scale seasonal and long term velocity dynamics of glaciers, however, remain an elusive goal.

The optical Landsat mission has a long history of monitoring and is thus useful for long term flow velocity analysis. Feature tracking algorithms applied on these Landsat images provide means of automatically calculating large scale velocities. Automated approaches for large scale analysis are difficult because of the lack of validation and good filtering techniques to deal with shadows, surface changes and clouds. Large scale temporal analysis is even harder because the errors arising in the flow velocity calculations are often large compared to the velocities due to the short periods between the images from which the velocities are calculated.

This research proposes, implements and tests a new method for automatically creating large scale velocity time series using the optical Landsat database and feature tracking in the Himalayas. Where normally velocity time series consists of consecutive single velocity fields, the novel method uses combinations of velocities to estimate these single velocities. This method is tested against results from single velocity fields for the Everest region and the Karakoram region. A sensitivity and parameter analysis provides the best parameter settings for the new method.

The result is a novel method that provides validation, robustness and acts as a filter for erroneous velocities. When possible, the new method increases or retains the number of results while increasing the precision. Furthermore, when the number of results is lower, it is due to filtering of erroneous velocities. The new method also provides an error indication which could be of use in future research. The main sources of errors: geo-location and precision of the feature tracking algorithm, are shown to have a large effect on the results and should be as small as possible. The magnitude of these errors make it difficult to measure seasonal changes in flow velocity for slow moving glaciers. Furthermore, Landsat 4-5 TM results are shown not to be useful for dense time series as the geo-location accuracy and artifacts from the TM scanner have a large effect on the precision of these results. The Landsat 8 velocity results, however, show good similarity to results from other researches and show promise for future research.

The sensitivity analysis of this research was incomplete as the complete algorithm was not optimized and future developments might focus on improving this. Also, it was found that the proposed method was very time intensive. A correct choice for the method of calculating flow velocities might drastically decrease the computational effort and improve applicability to large area research. In total, the new method is not only useful for use on the Landsat database, but could be used to improve any optical or radar flow velocity time series.

Contents

List of Figures

List of Tables

1 Introduction

Glaciers can be found everywhere in the world. From Patagonia to Greenland to the Himalayas. They serve as an important unknown factor in predictions of sea level rise[15]. Furthermore they are important for freshwater supply in many parts of the world [5].

Understanding dynamical and physical processes of glaciers is therefor valuable information in predicting the effects of glaciers in global and local climates and freshwater supplies. The velocity field of a glacier is one of these parameters that is of importance as it used as a boundary condition, validation and input for glacier modelling. In this research a new method for obtaining time series of glacier velocities using images obtained from the optical Landsat satellite mission is proposed and tested. It focuses and is tested on use in the Himalayas, but if successful could be extended for use on glaciers around the world.

1.1. Glaciers and the Himalayas

The Himalayas are the biggest freshwater storage in the world outside of Antarctica and Greenland. Many areas in the Himalayas are characterized by steep high altitude alpine terrain. The region is very large, with big differences in height, see figure 1.1. Many different local climates and ecosystems [41] thus exist which create areas with different seasonal and long-term variability in glacier dynamics. This spatial and temporal glacial dynamic variability is of interest for future research.

The global warming of the climate has continued retreat of glaciers in many mountainous regions. This is also true in many parts of the Himalayas. Himalayan glaciers and their melt water are very important for fresh water, food and power supply [13] [15] [50]. Studies show that there is no uniform response however of Himalayan glaciers to climate change[39][18]. Not every glacier thus reacts in the same way to climate forcing. Some glaciers gain mass (positive mass balance) while others lose mass (negative mass balance). Next to mass balance, glacial factors of study are withdrawal or advance, glacier velocity and thickness. While mass fluxes and glacier thickness are harder to determine, glacial velocity and glacial terminus withdrawal or advance are easier to measure. The availability of many different types of glaciers, climates and terrain makes the Himalayas suitable terrain for this research.

Glacier surface velocities are used in models that use the velocity as input for calculating mass balances and run-offs[39][24][27][12]. The thickness and volume of glaciers are again important parameters related to climate forcing and global warming [13], but also contain important information for policy makers. In negative mass balance regimes, ice fluxes will generally decrease to adjust. Mass balances have been strongly negative in large parts of the world for the last decades [22][51]. Changes to glacial velocity distributions and absolute changes in speed seem thus to be linked to changes in mass balance and run-off [33][49].

Because it is now possible to focus on comparing and analyzing glacier velocities within and between large regions we can improve knowledge about glacier dynamics and its relation to climate in

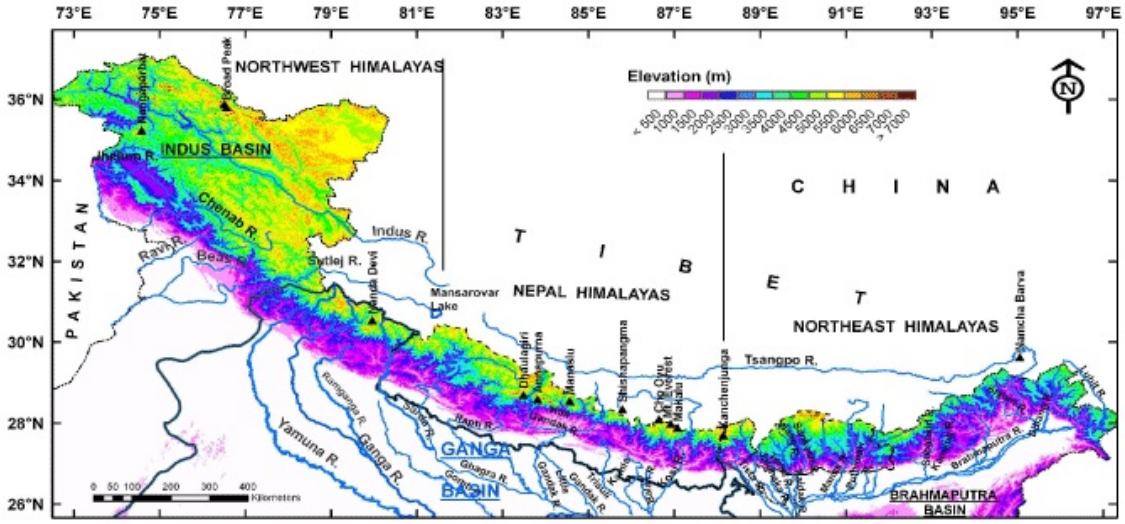


Figure 1.1: A digital elevation map (DEM) of the entire Himalayan mountain range [10]. It spans a large area and has a high elevation.

space and time. It is hypothesized and tested that a glacier having a negative mass balance will cause most glaciers to slow down and change the velocity distribution of a glacier. Generally and in five regions around the world this seems to be true. But this is only based on velocity differences between two epochs and generalized from small research areas to entire regions [18]. To gain a more complete view of relations between glacial mass balance and run-offs to glacial velocities, a bigger and temporally denser and larger network of glacial velocities should be available. Achieving this is the main goal of this glacial velocity research.

1.2. Obtaining glacial velocity fields

There are multiple techniques for obtaining glacial velocities based on GPS, UAV's and satellite imagery. For inexpensive and large scale glacial velocity fields, the use of satellites is beneficial. The techniques used for this are mainly based on offset feature tracking and InSar. InSar (Interferometric synthetic aperature radar) is based on phase wave interferometry between synthetic aperature radar (SAR) images. Displacement is calculated by radar signal phase changes between two epochs. Feature tracking calculates the displacement of a feature in the landscape (e.g rock, tree, crevasse) between two temporally spaced images. It can either be applied to optical satellite missions or radar satellite missions (SAR). Both have their advantages and drawbacks. This study will use the optical Landsat database. This is because it has a large database dating back to 1972 and good spatial and temporal coverage.

Satellite imaging in combination with advances in feature tracking algorithms has allowed for a global and repeated coverage of glacier velocity products and effectiveness of this method has been proven repeatedly [35][2][17]. A glacial velocity field is the calculated velocity for a large area of glaciers and surroundings.

Validation of glacial velocity fields from satellite imagery however is complicated. The results of feature tracking algorithms can contain faulty velocities which need to be filtered out, as in figure 1.2. Various techniques to do this exist and should filter out most faulty displacements. These techniques, however, also filter out correct matches and the filtering settings depend on the local area.

Up until now the physical nature of the optical satellite missions and the lack of usable satellite image pairs due to e.g clouds or shadows has not allowed for large area mapping with consistently short periods between acquisitions. Recent advances in feature tracking techniques and methods for pre-

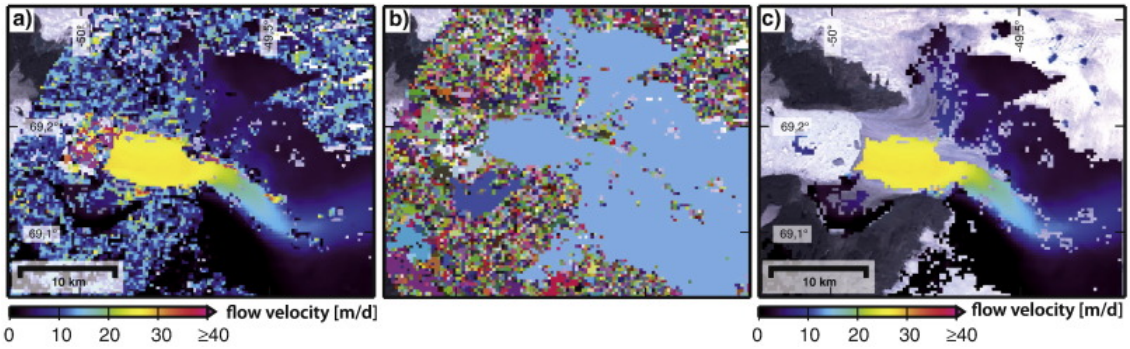


Figure 1.2: Glacier flow velocity field results near Greenland glacier Jakobshavn Isbrea derived from two Landsat images from 2000. (a) unfiltered velocity field. (b) filter application. (c) filtered flow velocity field. [35]

processing and post-processing have made it possible to automatically map large areas with robust results and less loss of information. Furthermore the introduction of the Landsat 8 mission opens up new opportunities for obtaining glacier velocities using feature tracking algorithms. [11][17][9]

1.3. Time series of glacial velocity fields

One of these new opportunities, which benefits from the addition of the new Landsat 8 satellite to the extensive Landsat database, is obtaining time series of glacier velocities. Opposed to previous Landsat missions, where the temporal coverage fluctuated significantly, the Landsat 8 mission has a very stable 16 days repeat coverage.

As seen in the previous sections it would be beneficial to glacial research if more dense time series of glacial velocities could be created using an automated approach. Recent research has focused on obtaining more glacial velocity fields in an automated manner. Time series of glacial velocities, however, are just started being researched. One example can be seen in figure 1.3. This research [11] uses a new software package (PyCorr) in combination with the high temporal resolution of the Landsat 8 mission to create time series of many Greenland glaciers.

Another recent research [35] also creates velocity time series for large areas from scenes acquired from the Landsat 7 mission, figure 1.4. This research uses spatial overlap between acquired scenes in combination with advanced filtering and pre-processing techniques to create a temporally denser, more robust time series.

Both researches only include time series of one Landsat missions from which only completely cloud free scenes are used. Other scenes are thus not used, while there may be useful data present. Except for spatial and temporal filtering, there is no validation of the data. These problems and restrictions could be solved by using the redundancy in combinations of scenes as in [9]. In this research a method is proposed to solve these problems and restrictions. This method combines multiple velocity fields to create a single velocity field. This could possibly provide an extra validation and increase the amount of data available after filtering. This research is partly executed at the University of Edinburgh where a processing algorithm is already partially developed. With this algorithm it is possible to create velocity fields from scenes of two epochs. This existing process is from now on called the "Base Method" (BM). The proposed method uses redundancy by combining (stacking) velocity fields as an extension of the BM and is thus called "SM" (Stacking Method) from now on.

1.4. Research questions and goals

The research question that fits the problems encountered in the introduction is:

Is it possible to make a glacial velocity time series from the complete optical Landsat mission for the Asian Highlands where the performance, i.e. the number of results and validity of results, is increased by

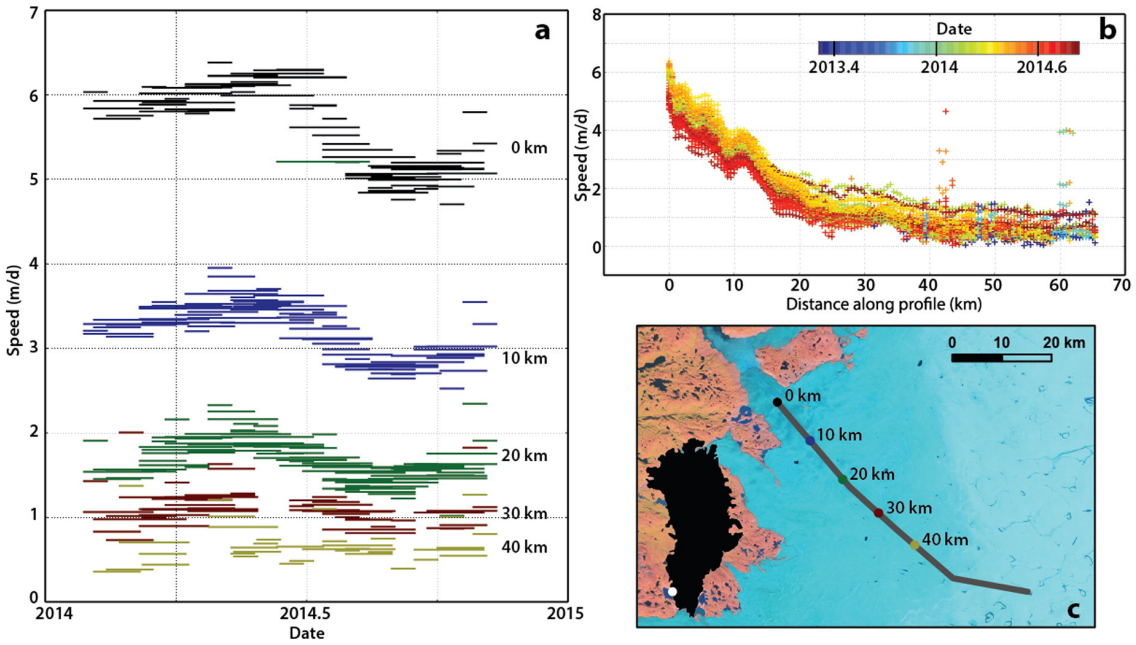


Figure 1.3: Landsat 8 velocity measurements for Kangiata Nunata Sermia glacier in Greenland. (a) Speed time series of one season for different points on the glacier. (b) Speed profile of glacier along profile line from (c) for different dates [11]

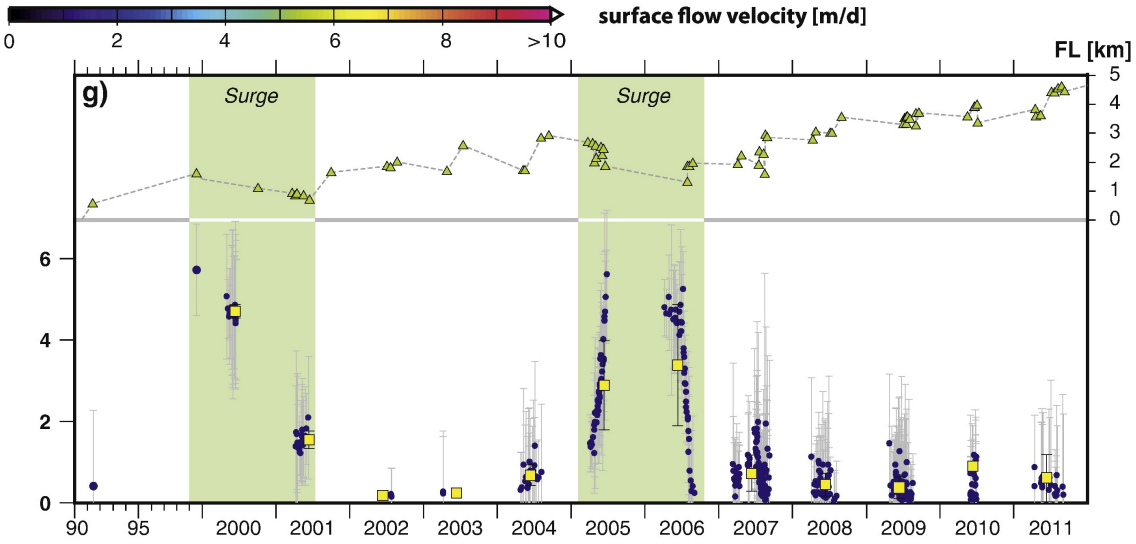


Figure 1.4: Flow velocity time series of a point on the Greenland Harald Moltke Brae glacier. Only includes Landsat 7 data. The two green areas are indicated as surge (sudden large flow) events. The data comes from multiple scenes. [35]

using combinations of velocity fields opposed to using single velocity fields?

Subquestions:

- **Using the current (BM) algorithm, could a precise time series be made spanning all Landsat missions?**
  The current setup allows glacier velocity fields for Landsat 1-7 scenes, but not Landsat 7 SLC-off or Landsat 8 scenes. Neither is it able to obtain inter-sensor velocity fields (for example combinations of Landsat 5 and 7 scenes). But is it useful for good performance to use all the data and

possible to combine different sensor scenes? Are there feature tracking methods that will allow for a more consistent time series?

- **What are the errors in the velocity fields, can they be quantified and how do they appear in the velocity fields?**
  Within what margin can we trust the results both from the current (BM) and new (SM) method. What could give errors and how are they handled?
- **What is the performance of a method and how can it be tested?** To test of the new method actually does better than the single pair methods, performance indicators should be devised and implemented that work for both methods and on time series. What are these indicators and how are they implemented?
- **How can combinations of image pairs be used to create time series and how should it improve performance?** What method could be used to use combinations of image pairs to estimate velocity fields, how can a time series be made using this method and why and how should it improve performance over single pair time series?
- **Which combination of feature tracking techniques, parameters and algorithm parameters gives the best result in the areas of research?** The existing algorithm and new method require setting parameters. These parameters should be chosen based on the research area, method and other factors. On which parameters and factors do the results depend and what is the best choice of parameters to create the best results?
- **What are the results of validation of the new method and application on the test areas?** Are results from the algorithm and the proposed method validated? Does the new method improve robustness of velocity time series and does it create valid results automatically for large areas?

To answer these questions, the setup of this research is as follows. More about Landsat and feature tracking can be found in chapter 2: 'Landsat & Feature Tracking'. The current processing setup, the proposed method and testing methods can be found in chapter 3: "Methods". Algorithm selection, algorithm issues and sensitivity analysis will be discussed in the 'Sensitivity analysis & Algorithm Selection' chapter, chapter 4. Results for two test areas in the Himalayas are presented in the 'Case Study' chapters, chapter 5 and 6. . Finally in chapter 7, the 'Conclusions & Recommendations' the results will be summarized and discussed and finally recommendations on improvements will be presented.

1.4.1. Research plan & setup

This research was conducted at the University of Edinburgh and consists of two parts. The first part is related to calculating the displacements and velocities of glaciers. This is the core of the algorithm. The second part is an extension of this core of the algorithm and is the creation of time series from these velocities or combinations of velocities. The core of the algorithm is outdated at the start of this research, it requires new feature tracking software, updating for Landsat 8 and changes to the pre-processing software. At the end of this research, the result should be an automated algorithm that can create dense unsupervised glacial velocity time series for large areas from Landsat images. To achieve this the following research goals are stated.

Research plan and goals:

- Adapt and update the core velocity algorithm to include Landsat 8 images.
- Update the core algorithm for creating velocity fields using a new feature tracking toolbox.
- Set up a processing environment en job setup for large scale processing
- Optimize the setup for a batch processing job of multiple Landsat sensors in terms of computational speed.
- Implement different time series algorithms.

- Testing of different methods
- Final implementation of selected method.

The theory for these goals is discussed in the next chapter.

2

Landsat & Feature tracking

The main factors of importance are the Landsat images used and the type of feature tracking algorithm. For a large part the quality of the results will depend on the type of data used, which is linked to the Landsat mission characteristics. Another important part is the feature tracking algorithm. The algorithm the foundation for all other processing steps. Significant details about both the Landsat data used and the feature tracking algorithm are therefor of importance for later analysis of the success and errors of the methods in this thesis. This chapter thus attempts to answer the questions: **Using the current algorithm, can a consistent time series be made spanning all Landsat missions?** and **What are the errors in the velocity fields, can they be quantified and how do they appear in the velocity fields?**.

2.1. Landsat

The Landsat program started in 1972 with Landsat 1 and is currently at Landsat 8, see figure 2.1. The extent of the the database (45 years) and large coverage makes that the Landsat mission is useful for the goal of this research.

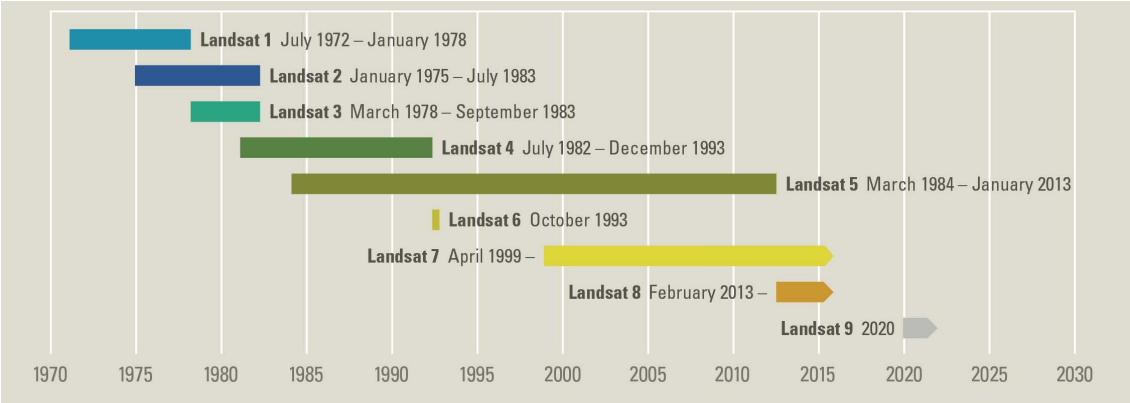


Figure 2.1: A timeline of all the Landsat missions including future missions. Landsat 7 and Landsat 8 are still functional [46].

There are however significant differences between the missions. The main difference between missions is the optical sensor. The sensor used for Landsat 1-3 was the Multi Spectral Scanner (MSS). Landsat 4-5 had an additional scanner called the Thematic Mapper (TM). Landsat 7 used an upgraded version of the TM called the Enhanced Thematic Mapper Plus (ETM+). The latest mission, Landsat 8, uses the Operational Land Imager (OLI) - Thermal InfraRed Sensor (TIRS). These different scanners use different spectral bands, have different radiometric performance, have different characteristics and even

work differently. The specifics of these scanners will be discussed in the section 2.1.2. Another difference between these missions is related to the orbits of the satellites. For example, the geolocation precision is higher for Landsat 8 than for Landsat 4 and the quantization of OLI-TIRS spectral bands is 12-bits, opposed to 8-bits in previous Landsat missions.The data is maintained by the U.S Geological Survey and can be downloaded free of charge as TIF images (also called scenes).

2.1.1. Landsat missions

From Landsat 4 onward the scenes are catalogued in the Worldwide Reference System 2 (WRS-2). This is a global notation system for Landsat data. It catalogues the earth and Landsat scenes into areas specified by a scene center and path and row numbers as shown in figure 2.2.

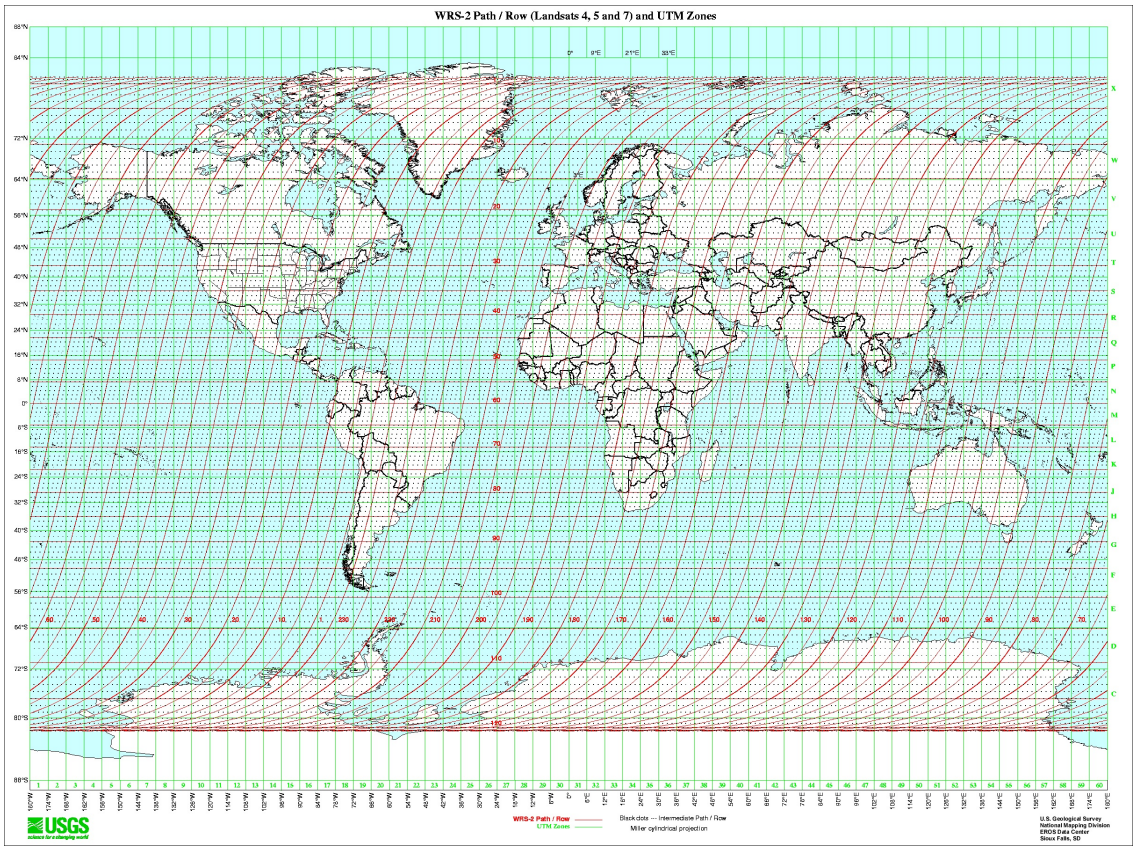


Figure 2.2: A visualization of the WRS-2 paths and rows plotted against the UTM zones [47].

The Landsat 4-8 missions are all operating in a polar, sun-synchronous orbit at an altitude of approximately 705 km. This means that the altitude and inclination (angle w.r.t equator) of the orbit are such that the satellite passes the same point of the surface at the same local solar time. The repeat coverage of these missions is 16 days. The satellites thus cross the same swath of surface every 16 days at the same local solar time. The Landsat 1-3 missions have a different reference system (WRS-1), a repeat cycle of 18 days instead of 16 days and slightly different orbit characteristics.

The main difference between these missions is the monitoring and calibration of the geometric orbit performance. Starting with Landsat 7, the geometric and radiometric performance has been monitored and calibrated precisely achieving improvements in geodetic as well as scanner accuracy and precision [23][42]. As obtaining correct displacement depends on the ability of the scanner to precisely geolocate the same pixel at different epochs as well as correcting if the pixels are misaligned, the geometric and radiometric performance is of significant importance. This will be discussed in detail in the next

sections.

2.1.2. Sensors

Both the MSS, TM scanner and ETM+ scanner are whisk-broom multi-spectral scanners. A whisk-broom scanner works with a mirror scanning across the ground track. This mirror deflects light into a single detector, scanning one pixel at a time,see figure 2.3. For a multi-spectral-scanner there are multiple spectral bands scanned simultaneously in this manner. The scene sizes collected by the TM are 185×172 km, while those of the ETM+ and the OLI-TIRS are 183×170 km, which is in accordance with WRS-2. The MSS obtained scene sizes are 185×185 km.

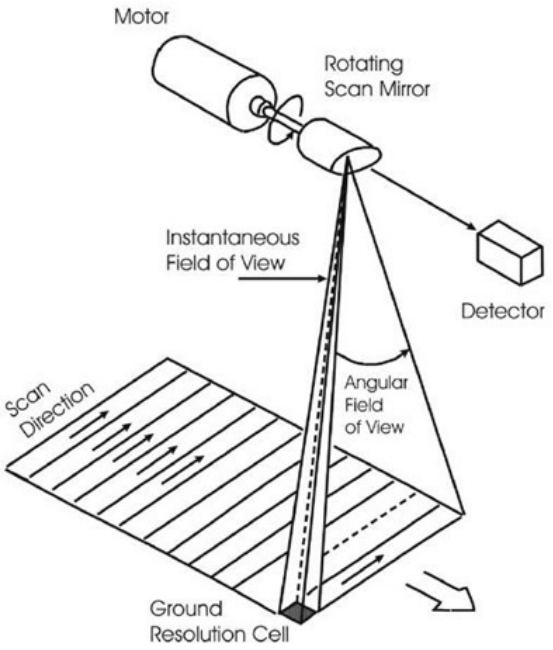


Figure 2.3: Schematic of workings of whisk broom scanner. Perpendicular to the the movement of the satellite a line is scanned 1 pixel at a time, as if sweeping a broom.

The OLI-TIRS scanner is a push-broom scanner. It uses a long linear array of detectors with thousand of detectors per spectral band to scan a complete line of pixels at once. There are thus no moving parts in this scanner and this setup also has advantages for the radiometric performance. Furthermore the quantization of the OLI-TIRS data is 12 bits instead of 8 bits for previous scanners.

The spectral bands also changed between scanners. A complete overview of all spectral bands per scanner is depicted in figure 2.4

A table of bands with resolutions (IFOV) for the MSS, TM, ETM+ and OLI-TIRS can be seen in figures 2.6 and 2.5. As of the ETM+ scanner a band with a resolution of 15×15m was added and the 120×120m TIR band was resampled to 60m resolution. The (IFOV) resolution of the MSS scanner is different from the other with 68×80 meter pixel sizes.

It is important to note that on 31 May 2003 the Scan Line Correcter (SLC) of the ETM+ scanner failed. This lead to stripes without data in the images, with a total of up to 25 % missing, as is showed in figure 2.7. The stripes are perpendicular to the direction of the satellite and wider near the edges of the scenes parallel to the movement of the satellite, while in the middle of the scene there are no data gaps. The images where the SLC failed are addressed as "SLC-off" .

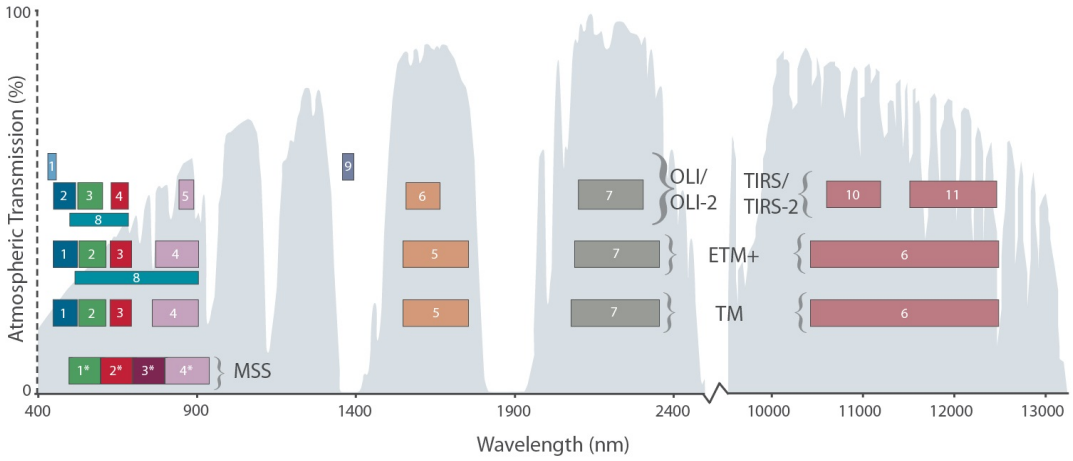


Figure 2.4: Visualization of spectral bands per Landsat sensor plotted as a function of wavelength against atmospheric transmission. [48]

Sensor	Band	Spectral range (micro m)		IFOV
MSS	4	0.50 ~ 0.60	green	80m
	5	0.60 ~ 0.70	red	80m
	6	0.70 ~ 0.80	near - IR	80m
	7	0.80 ~ 1.10	near - IR	80m
TM	1	0.45 ~ 0.52	blue	30m
	2	0.52 ~ 0.60	green	30m
	3	0.63 ~ 0.69	red	30m
	4	0.76 ~ 0.90	near - IR	30m
	5	1.55 ~ 1.75	interm - IR	30m
	6	10.40 ~ 12.50	thermal - IR	120m
	7	2.08 ~ 2.35	mid. - IR	30m

Figure 2.5: Table of spectral bands for the MSS and TM scanner with corresponding IFOV. [30]

2.1.3. Radiometric performance

Another important factor next to geometric performance and choice of spectral bands is the radiometric performance. As noted in the previous section, intensive performance monitoring started with Landsat 7 and the ETM+ sensor. The radiometric performance of the OLI-TIRS sensor is well characterized and is influenced by the type of sensor used as well as design of the optical system. One important difference between the TM, ETM+ and the OLI-TIRS is the quantization of the results. This is 8-bits for TM and ETM+, while being 12-bits for the OLI-TIRS. This means the images are being represented by arrays of intensity values ranging from 0-255 for 8-bits and a 0-4096 value range for 12-bits

Landsat-7 ETM+ Bands (μm)			Landsat-8 OLI and *TIRS* Bands (μm)		
			30 m Coastal/Aerosol	0.435 - 0.451	Band 1
Band 1	30 m Blue	0.441 - 0.514	30 m Blue	0.452 - 0.512	Band 2
Band 2	30 m Green	0.519 - 0.601	30 m Green	0.533 - 0.590	Band 3
Band 3	30 m Red	0.631 - 0.692	30 m Red	0.636 - 0.673	Band 4
Band 4	30 m NIR	0.772 - 0.898	30 m NIR	0.851 - 0.879	Band 5
Band 5	30 m SWIR-1	1.547 - 1.749	30 m SWIR-1	1.566 - 1.651	Band 6
Band 6	60 m TIR	10.31 - 12.36	*100 m TIR-1*	*10.60 – 11.19*	Band 10
			100 m TIR-2	*11.50 – 12.51*	Band 11
Band 7	30 m SWIR-2	2.064 - 2.345	30 m SWIR-2	2.107 - 2.294	Band 7
Band 8	15 m Pan	0.515 - 0.896	15 m Pan	0.503 - 0.676	Band 8
			30 m Cirrus	1.363 - 1.384	Band 9

Figure 2.6: Table of spectral bands for the ETM+ and OLI-TIRS scanner with corresponding resolutions. [48]

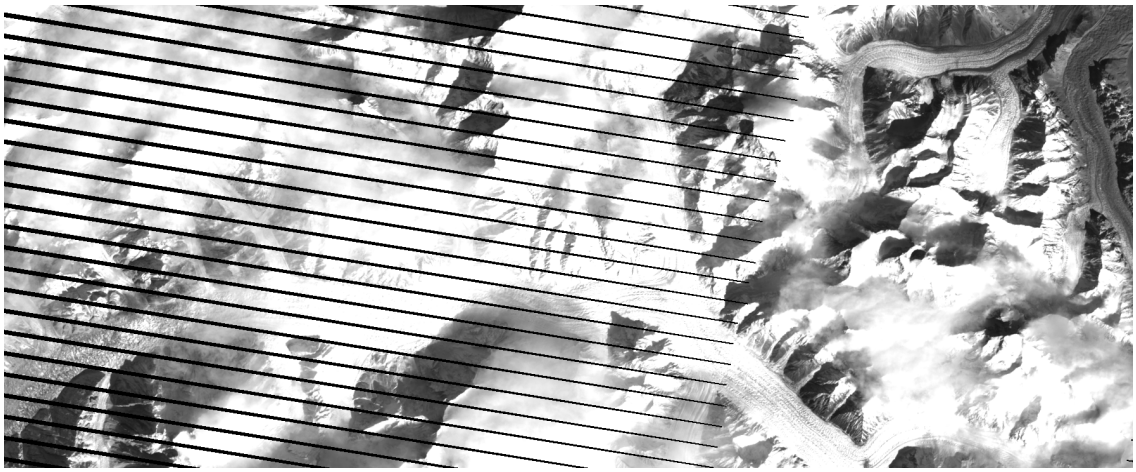


Figure 2.7: Segment of Landsat 7 SLC-off scene from 2004. The data gaps are wedge shaped and widest at the edges of the scene, while there are no data gaps in the center of the scene.

quantization. Other factors of radiometric performance are: noise, radiometric stability, pixel-to-pixel uniformity, artifacts, stray light and radiometric accuracy [29][26][25]. Another important factor is the system transfer function and accompanied point spread function. This point spread function (PSF) is an optical phenomenon that 'leaks' information from the pixel to neighboring pixels, see figure 2.8

Of these aspects only the noise (and accompanied signal-to-noise ratio) and radiometric stability have a large effect on feature tracking. The other aspects are not of importance as the effect is very small, or the effect is filtered out by the feature tracking method. The signal to noise ratio for OLI-TIRS and ETM+ nicely displays the improvement in radiometric performance between the two sensors. As shown in figure 2.9, the SNR of the OLI-TIRS sensor is 4-8 times larger than the ETM+ SNR. This has a positive effect on feature tracking as features are more accurately identified as such and not as noise.

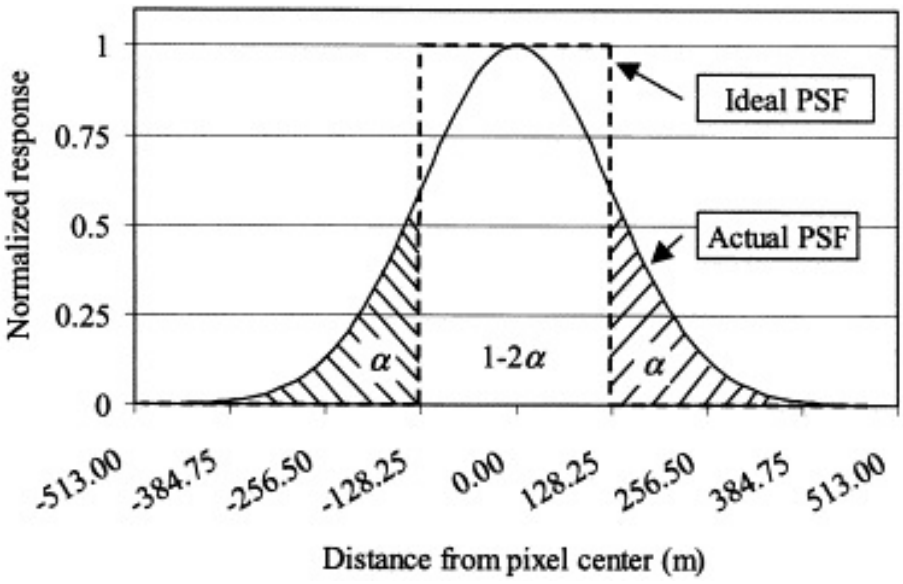


Figure 2.8: Example of point spread function. An ideal PSF would make that all radiance (response) is coming from the pixel itself. The actual PSF however includes radiance (response) from outside of the pixel. [20]

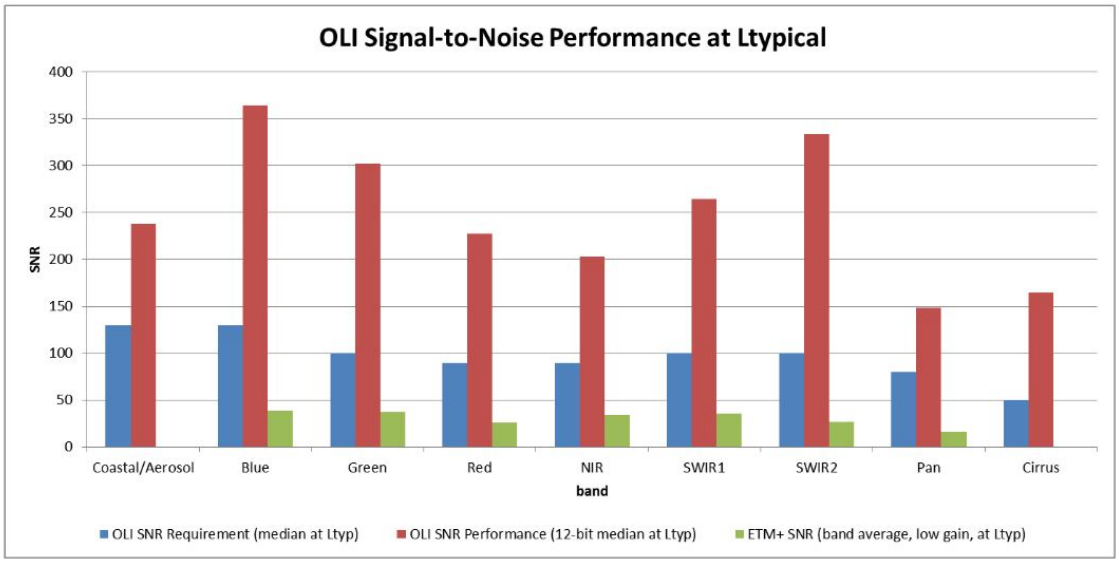


Figure 2.9: Signal-to-noise ratio of ETM+ and OLI sensors per spectral band at a typical radiance ($L_{typical}$) [48]

The SNR of the TM has been calculated in a different way, which makes it hard to compare the values.However, the results are in the same order of magnitude and it is suspected that the SNR would be lower or about the same for the TM and the ETM+ as the systems are very similar [19].

The second factor of importance is the radiometric stability. As this research is interested in temporal changes, the radiometric performance should be stable in between scenes. If this is not the case, artificial errors could be induced in the feature tracking algorithm. Studies have found results for the Landsat 5 TM, Landsat 7 ETM+ and Landsat 8 OLI. The Landsat 8 OLI radiometric stability is better than 0.3% over 60 days [29], and there seems to be no significant trend. The ETM+ stability is better than 2% over 12 years [26]. The stability of the Landsat 5 TM is about 5 – 10% [19], but corrections have been

implemented in the Landsat 5 TM scenes [7].

The final factor is the point spread function, PSF. As shown in figure 2.8, when the PSF is larger than a pixel, the intensity value of a pixel could change due to 'leaked' intensity from neighboring pixels. This might be a problem for feature tracking algorithms trying to characterize single pixels, like scale invariant feature tracking (SIFT). For larger areas of pixels the effect of the PSF is that it smooths the area (e.g decreases the variability) [20]. Effort has been made to characterize the system transfer function (STF) and accompanied PSF for the TM, ETM+ and OLI-TIRS [40] [43]. This shows that there is indeed significant leaking and smoothing due to the PSF in the TM and ETM+. Research on the effect of the PSF on quality of the repeat image feature tracking (RIFT) algorithm has however not been done. The expectation is that, because of the use of RIFT algorithm in this research which uses an area of multiple pixels and the use of several image enhancement techniques, this effect does not have a significant impact on the results. And that the effect is smaller for the Landsat 8 OLI-TIRS than the previous sensors due to the 12 bit quantization. These factors have also partly been examined for the MSS in [25]. The radiometric stability for MSS is about 15% over a three year period. The SNR and PSF have not been examined in detail for the MSS, but are believed to be worse than the TM.
There are also other effects inherent to the sensors, Landsat missions, geography and atmosphere which affect the quality and accuracy of the results. These will be discussed in the next section.

2.1.4. Quality and accuracy

The quality and accuracy of the Landsat scenes depends on multiple factors. The quality of the image depends for the most part on the atmosphere and illumination. Clouds obscuring the surface affect the quality in the most obvious way. Solar angle with respect to the surface is also a factor as it casts shadows on the surface obscuring features.

The accuracy of the Landsat data is expressed as the geometric performance. This includes e.g band-to-band registration accuracy, geodetic accuracy (geolocation accuracy) and image-to-image registration accuracy. Combined with orthorectification errors related to the angle of view and surface elevation this determines the accuracy of consistently locating the same point on the earths surface. Extensive research on quantification of these accuracy's has been conducted, especially for Landsat 7 and 8 [23][42].

As of 2016, however, the USGS started to manage the Landsat archive by reprocessing and structuring the Landsat images [44][46]. This approach identifies the Landsat images based on the quality parameters above and reprocesses the images based on this quality. The result is an arrangement into three categories and two tiers. The first category is L1TP, precise and terrain corrected (orthorectified) scenes, suitable for time series analysis. The radial root mean squared error of alignment of pixel data in this tier of data is < 12 relative root mean square (RRMSE), with scene specific RRMSE reported in the metadata file of the scene. The RRMSE is the root mean square error between the ground control points (GCP's) of the specific Landsat scene as compared to the average GCP's of all images of the same path and row. The other two tiers are L1GT, systematic terrain corrected scenes, and L1GS, systematic corrected scenes. These scenes do not have the geometric accuracy of the L1TP tier. This is because the reprocessing is based on registration with ground control points (GCP's). If too few or no GCP's can be used in a scene, for example due to clouds, the registration does not work and the scene is classified as an Tier 2 or L1GT image. The results of this is that every Landsat scene starting with the Landsat 4 mission from the same path and row has it's pixels aligned within the RRMSE provided in the metadata file.

A final important factor linked to the accuracy of aligning pixels between scenes has to do with the orthorectification.

2.1.5. Orthorectification

Orthorectification is linked to undulations on the earths surface and in the atmosphere and the angle of view of the satellite sensor. Changes in either of these affect the distance of which features are displayed (mapped) in the scenes, which would imitate displacement but is actually an error. Two images of the

same path and row taken from a different angle with respect to the nadir from a satellite could obtain a different position in the horizontal direction if the elevation of the surface is not well known, as is shown in figure 2.10.

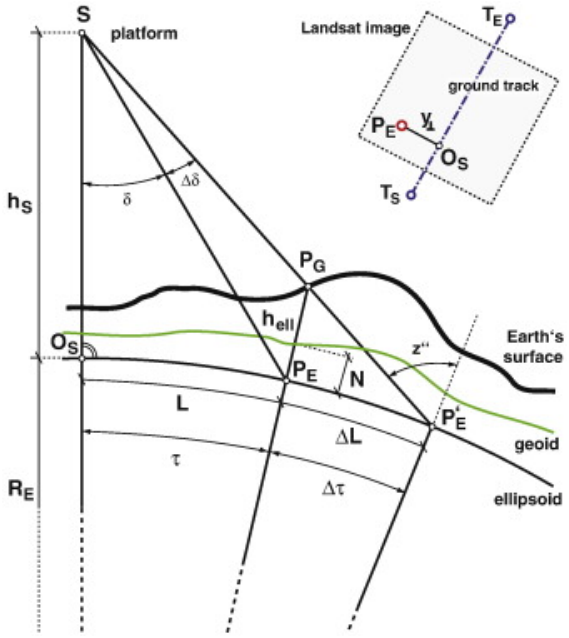


Figure 2.10: Orthorectification explained. An error in the digital elevation model (DEM) could map the same pixel displaced over a horizontal distance ΔL

A lot of research on using optical imagery to obtain glacial velocity fields focuses on this aspect as the errors could range up to several hundreds of meters [35] [38]. However, as in this research only images are used from the same path and row and changes of angle of view of the sensor are very small between scene acquisitions this error is assumed very small [18]. One problem with this is that it assumes that the surface height does not change between temporally spaced scenes. This is true for rockbed. When observing glaciers however, changes in glacier thickness occur due to seasonal changes. A change of 39 meters in elevation due to changes in glacier thickness between acquisitions can cause a shift up to 1/3 of a pixel for low angle satellite views [18]. Furthermore it also leads to surface decorrelation. This error is impossible to model as it requires precise knowledge of changes of the surface elevation. As discussed later on however, the method used in this research might filter out this error.

Most of the Landsat specifics discussed above are not area specific. To gain more insight useful for this research we take a look into Landsat scenes in the Himalayas.

2.1.6. Landsat scenes in the Himalayas

The overall number of scenes per Landsat mission for the Himalayan area has been mapped in figure 2.11 against time.

As can be seen, there is a data gap for the years 1980-1990. From 1990 onward Landsat 5 acquires more scenes using the TM, while there are hardly any Landsat 4 TM (LT4) scenes available. The number of Landsat 5 scenes increases around 2000. The amount of scenes from the Landsat 7 mission peaks in the first couple of years and has a steady number of acquisitions after around 2004. The most important conclusion from this figure 2.11 is that the temporal density is uneven and in some periods very sparse.

These Landsat images/scenes are the input for feature tracking algorithms. These feature tracking algorithms take two images of the same area with co-registered pixels of two different epochs and out-

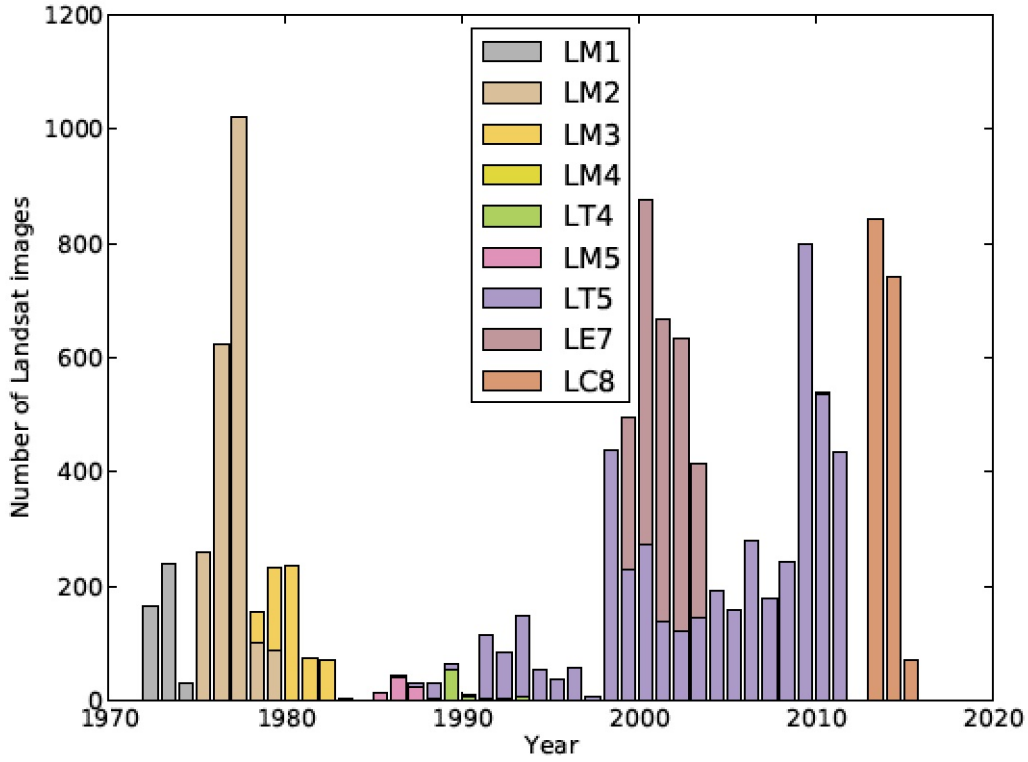


Figure 2.11: Number of Landsat scenes binned by Landsat mission as a function of time. There is a small data gap around 1985 and between LE7 and LC8. This data is from 2015 and more Landsat 8 scenes have been added since. [9]

put displacement fields in pixels. To do this there are several methods and feature tracking techniques clarified in the next section.

2.2. Feature tracking

Obtaining glacier velocities using satellite imagery is based on feature tracking algorithms. The principle behind feature tracking is simple and has been applied manually before the development of automated feature tracking algorithms. In feature tracking, features in images (e.g. distinguishable features like corners, edges) are identified in a combination of images. Application of this in glaciology relies on combinations of images that have aligned pixels, e.g the mapped location of overlapping pixels in two temporally spaced images is the same. Tracking the displacement between features (sets of pixels) in a combination of images is thus displacement due to moving objects, which could be clouds or glaciers. For example: identifying the same glacial crevasse in two images and calculating the distance (converted from pixels to meters) divided by the time between the images gives the velocity of the glacier at that crevasse .

The advance of computerized feature tracking started with the development of correlation algorithms and advances in digital imaging processing techniques [37]. In images where the resolution is high enough, features can be tracked. The two most common feature tracking algorithms are scale invariant feature tracking (SIFT) and repeat image feature tracking (RIFT). The most commonly used algorithm in glacial velocity research is the RIFT algorithm, which is a correlation algorithm.

2.2.1. RIFT

This correlation algorithm works as follows. First images are cropped, so that they have the same dimensions. Then a square subset of x by x pixels, called the reference chip, is chosen from one of the images which is called the 'master' image, see figure 2.12.

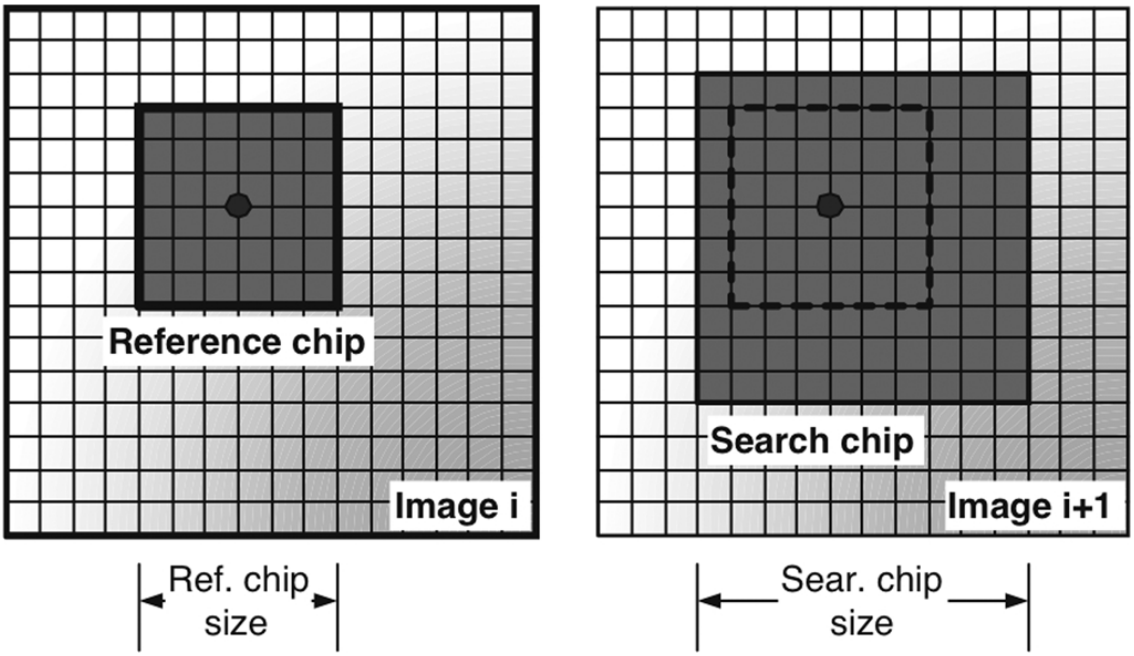


Figure 2.12: Feature tracking correlation algorithm chips. Image from [2]

This reference chip is then compared in the second second image (the 'slave' image) to all subsets of pixels in a search chip of y by y pixels, see figure 2.12. The comparison is based on a correlation between the two subsets of pixels. Thus, the reference chip moves over subsets of pixels within the search window. The result is a map of correlation intensities at the centroid of each reference chip position.The correlation of the centroid (i,j) in the search area with respect to the reference area is given by:

$$CC(i,j) = \frac{\sum_{k,l}[s(i+k,j+l)-\mu_s][r(k,l)-\mu_r]}{\sqrt{\sum_{k,l}[s(i+k,j+l)-\mu_s]^2 \sum_{k,l}[r(k,l)-\mu_r]^2}} \tag{2.1}$$

where (i,j) are pixel positions in the search chip, (k,l) pixel positions in the reference chip, r and s the values of the pixel in respectively the reference and search chip, μ_r and μ_s the average value of respectively the reference and search chip and $CC(i,j)$ the peak correlation coefficient of search window at pixel (i,j). The location of peak correlation in the 'slave' image with respect to the centroid of the reference chip in the 'master' image is the displacement of the feature between the two images. This method is called normalized cross-correlation or 'NCC'.

The cross-correlation function can also be operated in the Fourier domain [34]. Doing this, we get four more image matching methods which differ slightly. The first is the cross-correlation function of NCC equivalently applied in the Fourier domain, which is called 'CCF'. The cross-correlation function then becomes:

$$CC(i,j) = IFFT(F(u,v)G^*(u,v)) \tag{2.2}$$

where F is the Fast Fourier Transform(FFT) of the window with pixels u and v in the master image and G the FFT of the window with pixels u and v in the slave image and the * is the complex conjugate.IFFT stands for the Inverse Fast Fourier Transform

The advantage of techniques operated in the Fourier domain is that they have increased the processing speed. The drawback of CCF however, is that the normalization of the results is not computed. Differences in illumination (e.g intensity) between the two images or within a search area therefore can lead to mismatches[17]. One way to deal with this problem is to only use the phase information in the Fourier domain, also called 'phase correlation' or 'PC'. The biggest drawback of this method is that it experiences problems in areas with deformation, as the phase differences between images will not agree.

The two other methods are the same as CCF and PC, but operate on orientation images, thus called 'CCF-O' and 'PC-O'[14][16]. An orientation image of image f, call it f_0, is:

$$f_0(x,y) = sgn(\frac{\delta f(x,y)}{\delta x} + i\frac{\delta f(x,y)}{\delta y}) \tag{2.3}$$

where sgn is the signum function and x and y are pixel indices:

$$sgn(x) = \begin{cases} 0, & \text{if } |x| = 0 \\ \frac{x}{|x|} & \text{otherwise} \end{cases} \tag{2.4}$$

These orientation images split up the original image f into a real and complex part, where the real part is the intensity gradient in the x-direction and the complex part the intensity gradient in the y-direction. The advantage of this technique is that orientation correlation is illumination invariant [14] and can also be used on SLC-off Landsat images. Another advantage of OC is that the correlation is not affected in uniform areas. This means that the OC method is expected to have a better performance on snow covered glaciers. This also means that PC-O actually uses two kinds of normalization. One by using the orientation correlation and one by only using the phase. This could be a drawback as it removes too much of the original signal, while having the drawbacks of the PC as well.

RIFT algorithms commonly output three values: displacement in the x-direction, displacement in the y-direction and a signal-to-noise ratio (SNR). The SNR is the value of the peak correlation (e.g the maximum correlation found between the images) divided by the mean correlation. The mean correlation is the mean of all correlation values in the search window. The SNR is thus an indication of how certain the algorithm is that the displacement is correct. For example, when the feature tracking algorithm works on clouds, the SNR will be low because every template in the search window looks the same and thus the value of the peak correlation will be close to the average.

Success of these correlation algorithms thus depends on the type of terrain and the availability of recognizable features in the terrain, as well as the temporal baseline (time between scenes) between the master and slave image. A glacier is often not smooth and has debris on the surface like in figure 2.13, these are the recognizable features on the glacier.If however, for example, a glacier has been covered in snow between the master and slave image, it might be harder to recognize the same feature in both images. On the other hand, stationary features on glaciers may also exist, returning zero displacement results on glaciers. There are thus two kinds of errors induced by the RIFT algorithm. First of all, the exact location of the centroid is calculated by the RIFT algorithms with a maximum precision. Literature suggests the precision of centroid location within a pixel can be up to 1/10th - 1/20th of a pixel [17]. Second, correlation mismatches can be made. A correlation mismatch is a faulty tracked feature. Feature tracking applied on snowy patches on mountaintops with little features and no displacement for example could result in displacements, where there is actually none. Or, the other way around, in

calculating displacement between two scenes with a large temporal baseline, the features might have disappeared and the results might be a small displacement while there should be a large one.

Figure 2.13: Picture of the Baltoro glacier and surroundings. Notice the debris and the structures on the surface of the glacier,the features that are used in feature tracking.

To deal with faulty displacements several pre-processing, post-processing (filtering) and correlation algorithms have been developed. There are too many of these techniques to discuss here, the 'Methods' chapter will introduce the selected methods.

2.3. Summary

The Landsat database is suitable for glacial velocity monitoring because the missions have been operating for 45 years. Using the same path and row and repeat image feature tracking for the analysis in combination with the reprocessing of the Landsat scenes into a collections database greatly improves the precision and accuracy of geo-location. The increase in radiometric performance Of Landsat 8 OLI-TIRS improves the ability to recognize features in the images. The performance of OLI-TIRS Landsat 8 is better for RIFT glacial velocity fields [21]

The collections do not include Landsat 1-3 scenes. Together with poor expected radiometric performance, resolution and different orbit and mapping characteristics makes these scenes not very suitable for time series research. The time series would be best by using Landsat 4-8 L1TP scenes. Errors for these scenes are already given in the Landsat meta data and are assumed to only exist of random co-registration errors, when using the RIFT algorithm. The orthorectification error could have quite a large effect between epochs, but this is hard to include or model. Another systematic error source is the precision of the RIFT algorithm, which can be determined. Other errors due to radiometric performance, atmosphere or illumination will propagate through the RIFT algorithm and come out as faulty displacements. Pre-processing, filtering and post-processing have to be applied to remove these faulty displacements.

There are many spectral bands or combinations of bands to choose from. Also, the temporal spacing for older Landsat missions is not always equal to the repeat cycle. Combined with the failure of the SLC in the ETM+ scanner, this makes that temporal and spatial gaps in the data exist.

In the next chapter the functioning of an already implemented algorithm to obtain velocity fields is explained. Selection of spectral bands, RIFT parameters and types of RIFT algorithm are discussed. Furthermore an expansion of this algorithm for time series analysis is introduced and methods are proposed to deal with the temporal and spatial data gaps as well as faulty displacements with a minimum of user interaction. Finally the chapter discusses the errors that may arise in the velocity calculations.

3

Methods

As discussed in the previous sections, a complete processing setup, suitable for creating and analyzing time series of glacier velocities encompasses more than just selecting the correct Landsat scenes and applying a feature tracking algorithm. Multiple pre-processing techniques and Landsat mission specific adjustments have to be applied. Also, several choices of parameters impact the results of the algorithm. The algorithm to create the time series has to extend this setup and the outcome has to be post processed to filter out faulty displacements. Furthermore a method of testing these time series and glacier velocity field performances has to be set up. This chapter addresses this and attempts to answer the questions: **What is the performance of a method and how can it be tested?** and **How can combinations of velocity fields be used to create time series and how could it improve performance?** and also partially answers the question **Which combination of feature tracking techniques, parameters and algorithm parameters gives the best result in the areas of research?**.

The complete algorithm is depicted in figure 3.1.

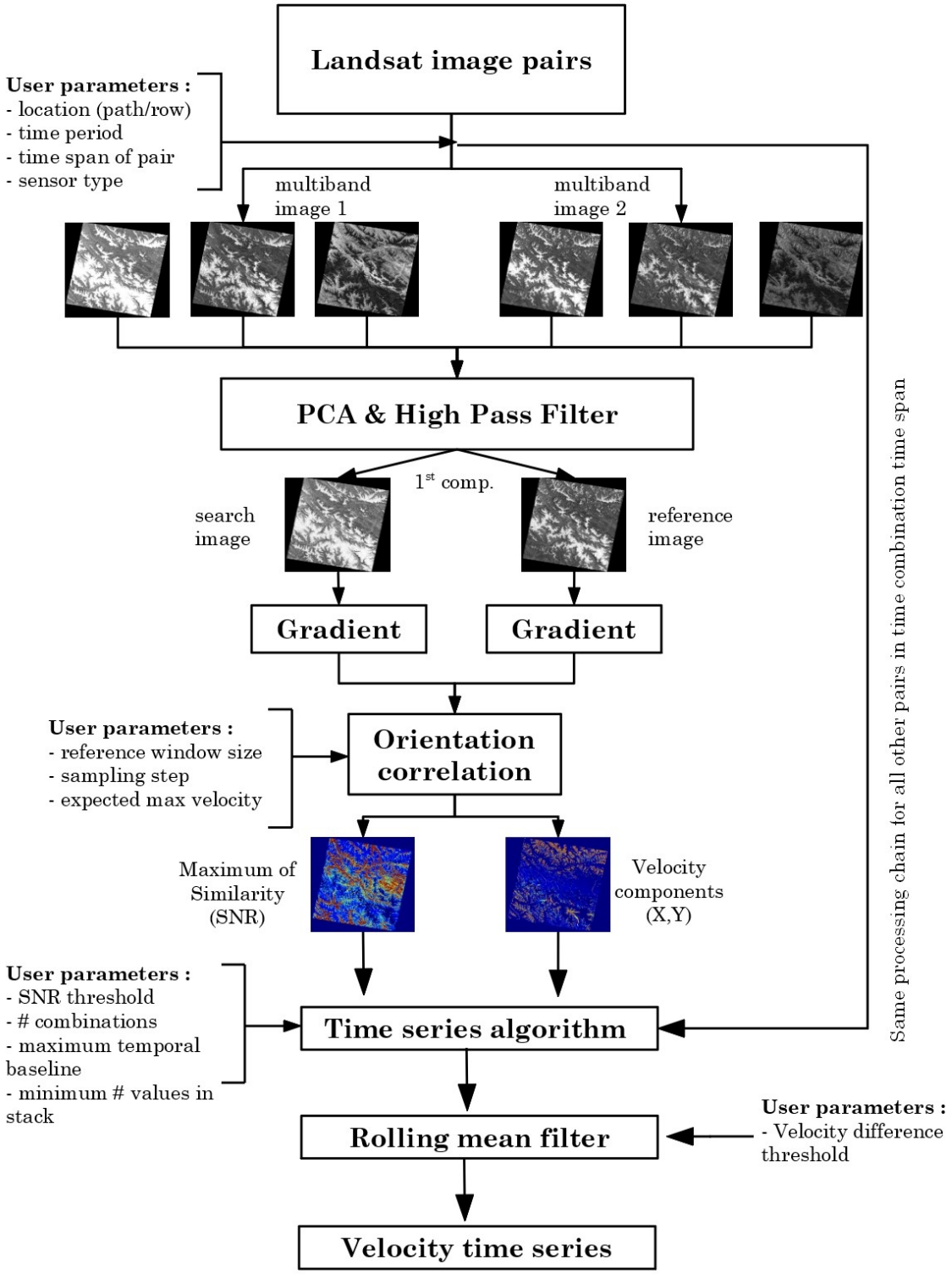


Figure 3.1: Final algorithm for obtaining time series of glacier velocities. This setup works for Landsat 4-8 scenes and can process image pairs from different paths and rows for a given time period with chosen time spans. The input parameters are stated per algorithm step and are summarized in table 3.3.

This algorithm consists of two main parts. The first part is the part of the algorithm which creates

the velocity fields. The second part is the proposed novel method for time series creation from these velocity fields. The core of the algorithm which calculates the velocity fields is discussed in the first part of this chapter "Calculating velocity fields". The proposed new methods for time series creation from these velocity fields, testing of the methods and the errors are discussed in the second part: "Novel method: Time series from redundant velocity fields".

3.1. Calculating velocity fields

3.1.1. Landsat scenes and data periods

To answer the research questions, it is necessary to take a look at different areas (e.g Landsat scenes). This is because of two reasons. First of all, the algorithm might fail in one region due to specific properties of the geography (steepness of the terrain, surging glaciers, snow cover, cloudiness of scene). Secondly, different areas in the Himalayas might experience different climate forcing and trends and thus the glaciers in different areas have different velocities and have different behaviour over time (different dynamics).

The areas considered in this research are shown in figure 3.2 and table 3.1.

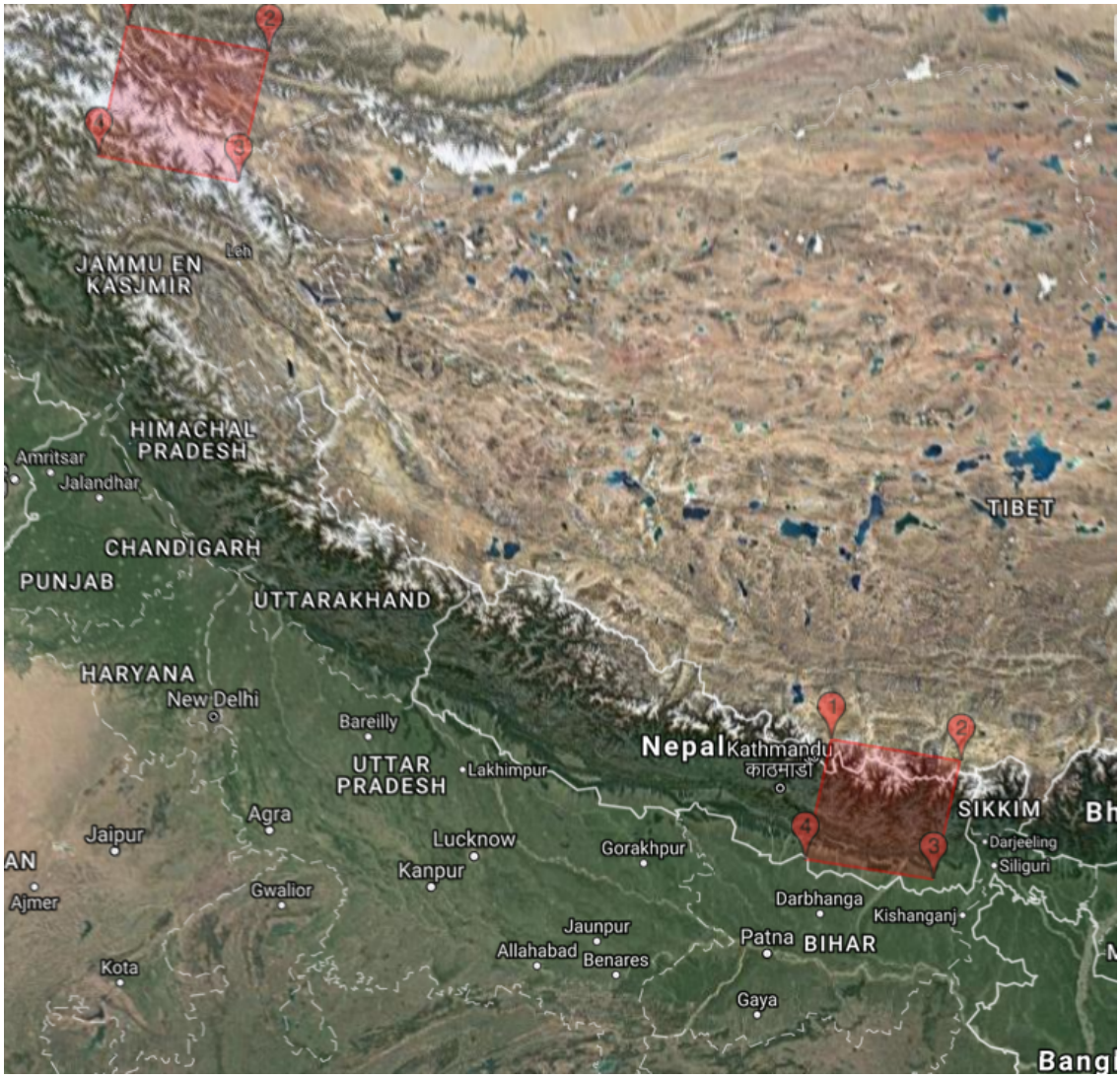


Figure 3.2: The areas of research. Top left is the Karakorum region. Bottom right is the Everest region.

Region	WRS-2 path	WRS-2 row	Center Lat	Center Lon
Everest	140	41	27.432	86.796
Karakoram	148	35	36.044	76.754

Table 3.1: Research areas in the Himalayas. Name of region with WRS-2 information and latitude and longitude of center of the scenes.

Each of these areas corresponds to a WRS-2 defined area. For each of these areas all available Tier 1 (L1TP) TM, ETM+ and OLI-TIRS data is selected. That is, all images starting from Landsat 4. The L1TP tier contains the largest amount of images. For example, only 6 images out of 120 are not L1TP images for TM scenes in the Karakoram area. Only L1TP is selected, because it provides quantitative information about the precision of the co-registration of images and is "suitable for time series analysis" [46]. Most scenes categorized as Tier 2 also have a very high percentage of cloud cover, making them less useful for this research.

All scenes which fit these criteria are then subdivided by sensor. The main reason for this subdivision is that the correlation between scenes observed with different sensor types is not possible in the current setup of the algorithm. The different spectral bands with different resolution for different sensors is another reason feature tracking between sensors is not implemented. These subdivisions, which are L1TP Landsat scenes for a specific sensor, are the input of the algorithm. This selection of scenes per sensor type is then first pre-processed to enhance features before the feature tracking is applied.

3.1.2. Selection of bands and importing data

Every sensor scene consists of multiple spectral bands. It is important to select the correct spectral band or combination of spectral bands. Earlier research from [17] and [9] suggest a combination of spectral bands 4 and 5 for Landsat 4-5 and Landsat 7 scenes and spectral band 8 for Landsat 8. Although the resolution of band 8 for Landsat 7 is higher than the resolution of bands 4-5, the results are better for bands 4-5. This is because the combination of spectral bands of bands 4-5 work well for icy and snowy areas. These results are based on the total number of correct velocities in a scene.

The selected bands from each scene of each epoch are then cropped to the size of their respective WRS-2 areas. If needed, a glacier mask can be created from a glacier GLIMS shapefile, indicating the positions of glaciers in the scene. This glacier mask not only includes flowing glaciers, but also slopes and peaks of nearby mountains or areas that have a constant ice or snow cover. Another option is to also co-register the scenes. As the selected scenes are assumed well co-registered and the expected error between bands of a scene is very small, these steps are skipped. Because of the same reason only one ice mask is created for all sensor images from a path and row. Two pre-processing steps then follow: first, principal component analysis is applied to the combination of spectral bands of a scene if multiple bands are selected. Afterwards, a high pass filter is applied to the images.

3.1.3. Spatial high pass filter & Principal component analysis

The first image enhancement technique is principal component analysis (PCA). PCA is a technique that sorts data on variability and splits the data based on this variability under the constraint that the separate data sets are orthogonal to each other and maximize the variance [37]. In glacial research it is used to combine the interesting features from multiple spectral bands into one image. The result is a new image as in figure 3.3, which uses multiple spectral bands.

In figure 3.3 the result is clear. The color scales are identical in both images, but the contrast between features on the glacier is more enunciated in the PCA enhanced image. Thus the contrast between features on glaciers increases when applying PCA which is beneficial for the correlation step.

After the PCA, a spatial high pass filter is applied to the images. The high pass filter should filter out low frequency components in the image that relate to the geography and quasi-stationary surface changes, while enhancing high frequency components corresponding to small scale features that are advected with the ice flow (like crevasses or rocks) and are due to basal topography, see figure 3.4.

The used filter is a Gaussian high pass filter. The input parameter is the size of the kernel in meters.

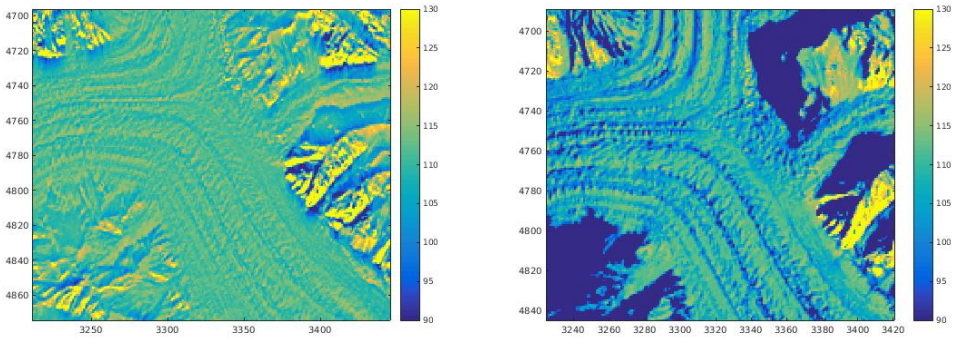


Figure 3.3: Extracts of the Baltoro glacier in the Karakoram region. Results from PCA step applied to the same Landsat 7 scene on band 4 and band 5.The color bars are 8 bit intensity values. Figure (a): no PCA applied, spectral band 4. Figure (b): PCA result, spectral bands 4 and 5.

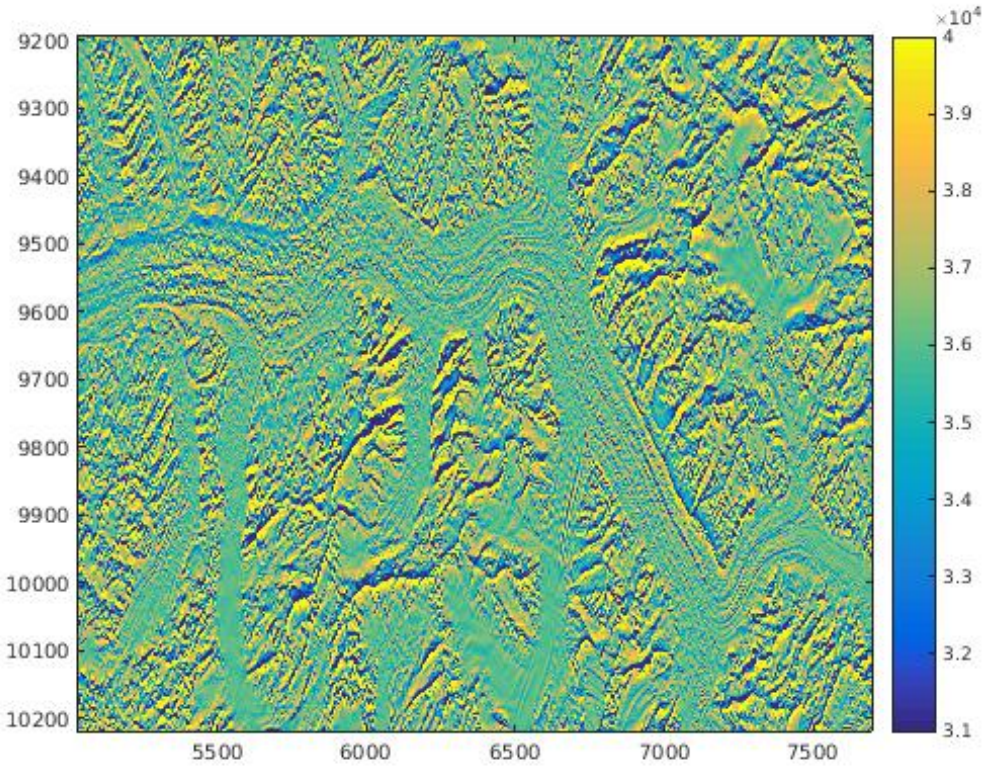


Figure 3.4: The Baltoro glacier in the Karakoram region. Results from the high pass filtering step applied to a Landsat 8 scene on band 8.

A kernel size that is too large might reduce the amount of recognizable features to track, while a kernel that is too small might have little effect on enhancing traceable features. A high pass filter kernel with a size of 1000 meters (33 pixels for 30 meter resolution, 66 pixels for 15 meter resolution) is selected. The results from the pre-processing steps are two images which have been enhanced by one or both of these techniques and can be used as an input to feature tracking algorithm.

3.1.4. Feature tracking: Imgraft

There are several commercial and non-commercial feature tracking packages available. The package of choice is the Imgraft (Image Georectification and Feature Tracking Toolbox) package [28]. The choice for this package is based on personal communication, which concluded that many other packages seemed to offset the centroid of the reference chip in the search image to integer pixels. Imgraft does not show similar behaviour. Together with the availability of multiple correlation methods this package is suitable. The chosen and implemented correlation method is orientation correlation (OC) based on research from [17]. This method gives the best results in terms of number of assumed correct results on glaciers in the Karakorum region.

There are two inputs and three parameters linked to the feature tracking algorithm and it outputs three results. The inputs are:

- Master image
- Slave image

And the parameters are:

- Size of search window
- Size of reference window
- Window Spacing

The outputs are arrays (which are essentially images) of a size $(M \times N)/W$, where $(M \times N)$ is the original size of the input image and W is the Window Spacing. These outputs are: a displacement in pixels in the x and y-direction as well as a signal-to-noise ratio. A velocity field is calculated from these displacements by:

$$v = r * (365/T) * \sqrt{(dx)^2 + (dy)^2} \tag{3.1}$$

with dx and dy the displacements in pixels in the x and y-direction respectively, r the resolution of a pixel, T the temporal baselines between the two input images in days and v the speed in m/yr.

The selection of the values of the input parameters depends on the area and can be evaluated by looking at the indicators of performance. Research and testing by [17][9] provides a reference value for the Karakorum area as well as an introduction into the effects of the choice on the results.

A reference window that is too small will only correlate noise and not features. A window that is too large will generalize too much. Local adaptive reference windows have been proposed and developed [8], but are computationally expensive. The size of the search window is automated based on expected velocities of the glaciers and the time span between the images. If the size of the search window is too small, the correct correlation might not be found. A search window that is too large takes more computational time and increases the chance of finding false matches. Finally the window spacing parameter is chosen to be half the size of the reference window. This creates overlap in neighboring results, while minimizing computational time. The chosen parameters for each region are shown in table 3.2 and based on research from [9]. The maximum velocity and size of the reference window (and thus spacing) state two options. Both options will be tested and the best parameter combinations will be selected and used for all testing and case studies.

From table 3.2 we can see that there are four parameters of such importance that they can alter the quality of the results: the size of the high pass filter, the size of the reference window, the choice of bands and the maximum velocity. The choice of type of correlation is also of influence. The setup has already been extended for the use of Landsat 8 and the use of orientation correlation at this step. Before the proposed time series method is explained however, the next section proposes the method for improving results from Landsat 7 SLC-off scenes.

Region	Sensor	band(s)	high pass filter (pixels)	reference window (pixels)	window spacing (pixels)	maximum velocity (m/yr)
Karakorum	LT5	4,5	1000 meter	12 or 16	6 or 8	500 or 800
Karakorum	LE7	4,5	1000 meter	12 or 16	6 or 8	500 or 800
Karakorum	LC8	8	1000 meter	12 or 16	6 or 8	500 or 800
Everest	LT5	4,5	1000 meter	12 or 16	6 or 8	500 or 800
Everest	LE7	4,5	1000 meter	12 or 16	6 or 8	500 or 800
Everest	LC8	8	1000 meter	12 or 16	6 or 8	500 or 800

Table 3.2: Imgraft parameter settings for test areas and sensor types. 12 or 16 means that these results will both be tested and compared in the results.

3.1.5. SLC-off images

Since part of the Landsat 7 scenes miss data because of the failed scan line corrector, analyzing these images is another challenge. There are several solutions available. The first is to mosaic overlapping Landsat scenes to fill these gaps. These Landsat scene gap filled products are provided by several parties, but have their limitations on small scale features[45]. It is therefore deemed unfit for use in time series analysis. The second solution is based on the correlation method used for the feature tracking. The SLC-off data is provided with a mask that indicate the gaps. The proposed solution is to fill these gaps with white noise.

The correlation function of white noise is a constant value of zero. Thus if a correlation is attempted in a data gap area, no significant correlation peak should be returned and thus the results should be easy to filter out. This solution thus does not solve obtaining data for the data free areas, but should make it possible to get a result for the rest reducing errors caused by mismatches due to the data gaps.

Having covered the steps for creating velocity fields, we take a look at the proposed algorithms for making time series using combinations of these velocity fields. It is discussed why this method might increase performance, what variations of the proposed method are tested, what parameters are of importance, how to assess performance differences for time series created with different methods and finally what the expected errors are and how they will affect the results.

3.2. Novel method: Time series from combinations of velocity fields

3.2.1. Proposed method

It was introduced that some of the problems of automatically creating time series using the Landsat data are validation, the sparseness of the data (there is no new image every repeat cycle for older Landsat missions) and the small temporal baselines. To improve this, a new method based on [6] and [9] is proposed that uses combinations of velocity fields to validate results and improve the amount of data available (coverage) for the time series. Opposed to for example, [35] and [11], where for every velocity at an epoch only one velocity field is used, the proposed method uses multiple combinations of temporally spaced images to provide validation and increase the amount of data (coverage). How the proposed method might improve coverage, increase performance and affect the error budget is discussed in sections 'Indicators of performance' and 'Precision & error budget'. The core of the proposed method is depicted in figure 3.5 and 3.6.

The method is based on expressing displacements between two epochs (in the x or y direction) as not only the actual displacement between epoch 1 and epoch 2, but also combinations of displacements. For example: to obtain the displacement between images pairs at t_6 and t_7, d_{6-7}, we can use d_{6-7} directly, but also combinations of other pairs. For example $d_{6-7} = d_{6-8} - d_{7-8}$. This is shown in figures 3.5 and 3.6. As the output of the feature tracking is a displacement between two epochs, i.e images, this combining of displacements can also be used for glacier velocity fields.

We can rewrite these combinations as:

$$d_{i-j} = d_{i-\delta m} - d_{j-\delta m} \tag{3.2}$$

when δm refers to a future epoch, or

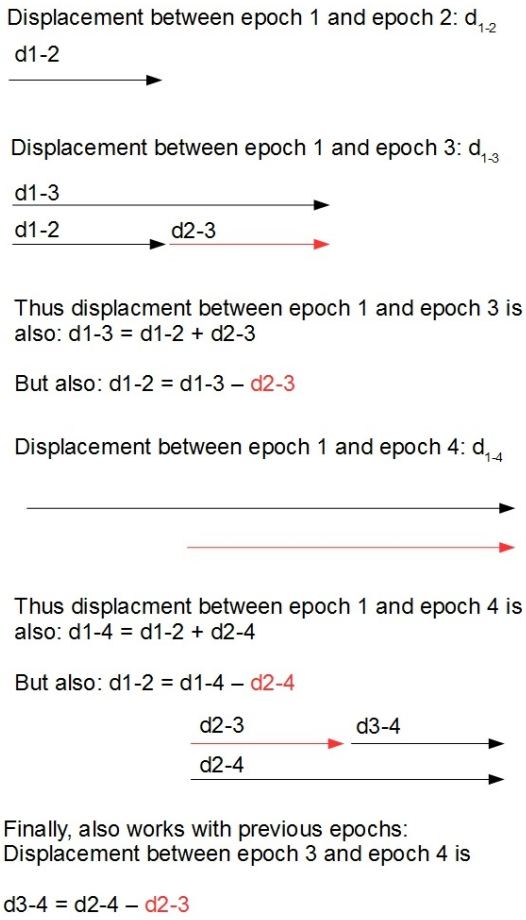


Figure 3.5: The displacements between two epochs can be expressed in multiple ways.

$$d_{i-j} = d_{\delta m-j} - d_{\delta m-i} \tag{3.3}$$

where δm refers to a previous epoch. δm is the number of indices of difference with the original velocity pair indices, d is the displacement and i and j are indices of epochs of the image.This means that if there are N displacement image combinations, the total amount of velocity fields will be $N + 2\delta m - \frac{\delta m(\delta m-1)}{2}$. The result is a series (stack) of δm displacements at a location and epoch.

As these displacements are theoretically all the same displacement, the next step would be to estimate the true displacement from this stack of displacements. This could be done using a mean estimator, but a median estimator is more robust against outliers. In [9] it was chosen to use the more robust median estimator for estimating the velocity value from temporal and spatial distributions. The estimator used in this method therefore will also be the median.

Based on this principle adisplacement estimate can be made consisting from δm combinations for every epoch. But the data is divided into sets per sensor type and at the beginning and ends of a set only forward or backward combinations can be used. To use the data in an efficient way the time series is created as shown in figure 3.7.

The first displacement estimate of a time series is made by using the first δm combinations. The second estimate made by using $\delta m - 1$ forward combinations and 1 backward combination and so on, until the middle of the series. In the middle of the series, when there are enough forward as well as

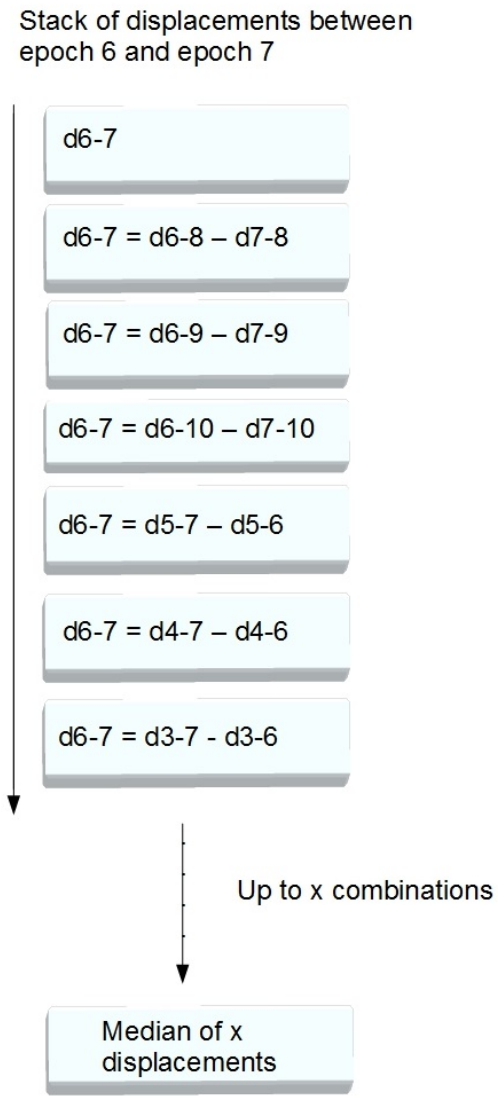


Figure 3.6: Combining combinations of displacements creates a stack of displacements from which a true displacement can be estimated.

backward combinations available the x combinations will be made up of $\delta m/2$ forward combinations as well as $\delta m/2$ backward combinations. At the end of the series the number of forward combinations decreases until the final epoch consists of δm backward combinations. This has 2 advantages: the temporal baseline is minimized and the number of velocity fields that have to be calculated are also minimized. Because in older sets of, for example Landsat 5 scenes, the temporal baseline is not always 16 days (the repeat cycle), when only forward combinations are chosen the combinations might consists of very long temporal baselines. To minimize this the combinations are split up into forward and backward combinations and a maximum temporal baseline of two years is set. In $d1-2 = d1-29 - d2-30$ for example, d1-29 is only used if the time difference (temporal baseline) is smaller then 2 years.

The functioning of the complete algorithm as shown in figure 3.1can be summarized as:

1. A list of pairs to be processed is created following the method in section 3.7 based on: time span,

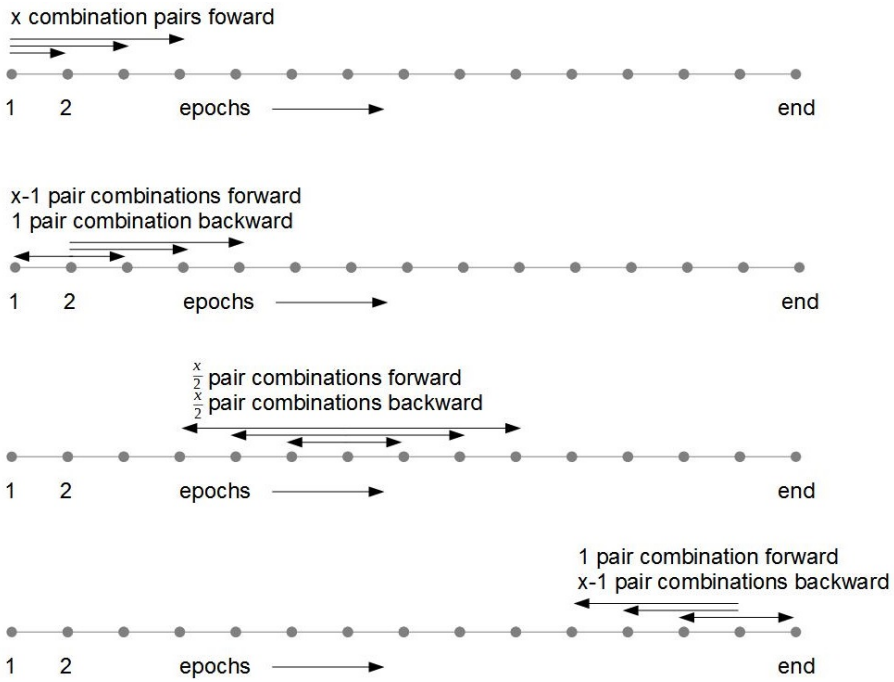


Figure 3.7: Example of stack of multiple displacements expressed by different combinations . The median of this stack of displacements is a estimate of the true displacement

Landsat mission, the WRS-2 path and row of the scenes (multiple allowed), minimum and maximum allowed temporal baselines and number of δm combinations (size of stack per epoch).

2. Every pair in this list is processed following the procedure of the core method. This is done in parallel. The inputs are: the list of pairs, selected bands, size of Gaussian filter, size of the reference window and the expected maximum velocity.

3. The output is a complete series of redundant displacement pairs in the x and y direction and a SNR field for a sensor type. The algorithm now iterates over all of the base pairs and stacks all combinations per base pair and direction, while filtering using the SNR. For example: it selects base pair dx_{1-2}, filters this on a specified SNR threshold. It then makes combinations after filtering the single velocity pairs on the SNR threshold. Thus, subtracting dx_{2-3} from dx_{1-3} after filtering both based on the same SNR threshold. The result is a stack of SNR-filtered displacements in the x and y direction, as showed in figure 3.8. The median of these stacks is estimated and a speed and velocity direction calculated.

4. The result is a series of velocity fields for every epoch, which is the time series. These can be post-processed by applying spatial and temporal outlier filters, which is covered in the section "Post processing and filtering".

There will be three variations of this method which will be tested, these are all variations of the stacking method (SM). The best method will be selected and applied to the test areas, while the performance will be compared to the base method.These three variations of the SM method are:

1. Filtering each displacement field before creating combinations and taking the median of only the same pixel displacements in the stack. Thus, before creating a displacement field $d1-2 = d1-4-d2-4$, $d1-4$ and $d2-4$ are filtered based on SNR. Furthermore the number of displacements to calculate the median has to be 5 or more values. This sub-method will be called "Single Pixel Stacking Method' (SPSM).

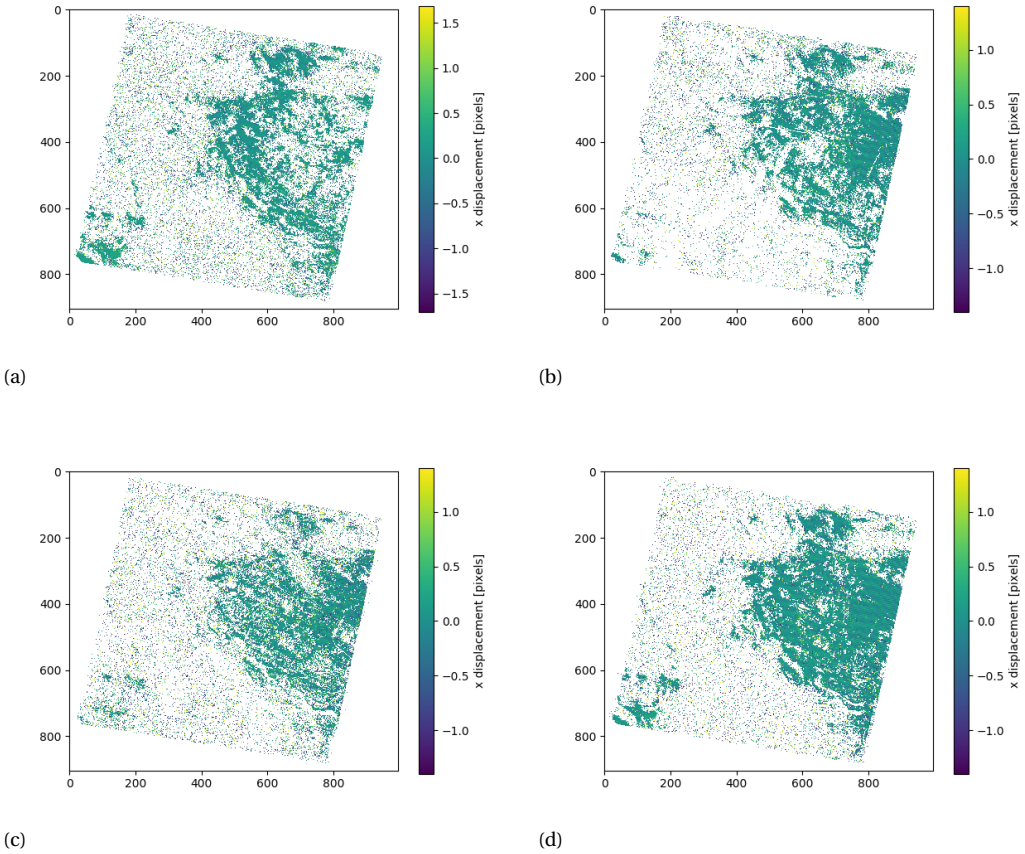


Figure 3.8: Selection of displacement fields in the x direction from a stack all representing the same displacement fields, filtered with a SNR threshold of 2. The temporal baseline of the base pair image (A) is 32 days. Images B, C and D the same velocity fields, but created by combinations: B = d1-8 - d2-8, C = d1-14 - d2-14 and D = d1-23 - d2-23

2. This method is the same as method number 1, except that there is an extra filtering step next to the SNR. The second filter is a spatial filter that removes all displacements without any spatial neighbors. This creates an extra form of validation and is expected to filter out outliers that were not filtered out by the SNR. This sub-method will be called "Spatial Filtered Stacking Method" (SFSM).

3. This third method differs from 1 because the median is not only calculated from the same pixel locations in the stack of redundant displacements, but also the 8 pixels surrounding all the displacements. This method applies some spatial validation as well, furthermore the median is calculated from much more values than the previous 2 methods. Therefore the threshold for the number of values before the median is calculated is set at 15. This sub-method will be called "Nearest Neighbors Stacking Method" (NNSM).

The best method of these three variations will be selected and used to test against the base method. The next section will discuss how this will be tested and what the performance is. Another important feature of this testing is validation. As noted in [17], validation is difficult. We distinguish two types of validation. The first one is validation by comparison with velocity fields from other methods. The second one is validation by testing temporal and spatial consistency. Multiple validation methods have been proposed and applied, based on temporal and spatial validation. This will be discussed in the " Post-processing" section. Finally the method of external validation is discussed and the errors are analyzed.

3.2.2. Indicators of performance and performance test

The previous section explained the algorithm and the proposed sub-methods for testing. But why might this new method improve results and how can the performance of these methods be tested? It was introduced that scenes are not used when there are clouds in the scenes and that using only the SNR and post-processing techniques often filters out valid points. This means that valid velocities might be lost. Furthermore if no validation data from other sources is available, the only validation solution is to check for temporal and spatial consistency. This might give problems in areas with surging glaciers or on edges of glaciers. Finally, assumed correct values that are actually false pose another problem. The proposed methods might overcome this by: not falsely filtering out correct values and the availability of a measure of consistency between velocity values at a data point. It must be made clear that there is a limit to the performance of this new method. In areas with clouds, shadows or low contrast results are unwanted or impossible and the proposed method should not "create" results for these areas. The main goal is to improve reliability of the results and possibly increase the coverage such that large scale automated time series analysis for glaciers is more feasible.

Two tests will be performed: the first is testing and comparison of the different sub-methods. The best method is selected, its functioning studied and tested against the base method in case studies of the two areas. The testing is based on the indicators of performance.

Indicators of performance

Several indicators of performance that are applicable to both the base method and the proposed SM methods are selected based on the stated goals.

- The success rate (SR): The ratio of the amount of assumed correct velocities on the glaciers over the total number of available velocities on glacier. The SR is expressed per velocity field and the total number of velocities is the total number of pixels on the glacier mask.

- The median absolute deviation (MAD) on stable areas. The velocity on stable ground should be zero if the algorithm were perfect. The precision around the stable velocity (e.g MAD) is thus an indicator of how precise the algorithm is. The MAD is a robust version of the standard deviation and estimates the standard deviation as $\hat{\sigma} = 1.483 * MAD$. Whenever the MAD is used, this estimate of the standard deviation is used.

- The MAD between the displacements in the stack used to estimate the true velocity: the dispersion. In the SM methods, the result is a stack of velocities in the x and y direction. The MAD between the values used to calculate the median is a measure of how well the estimated true velocity approaches the true velocity.

- Temporal success rate (TSR): The total number of assumed correct velocities for a single pixel, but for all epochs. Thus, where the SR indicates the spatial success, the TSR indicates the temporal success.

- Smoothness of the time series. It is expected that the velocities do not vary much between successive velocities. Smoothness can be determined in different ways, but is tricky in this application as the time series can be very discontinuous (e.g many points might not have a velocity). The smoothness is thus calculated for a time series of a pixel and the 8 neighbors. It is calculated by taking the one lag auto-correlation function. This is expressed as a value between -1 and 1, closer to -1 being more random and closer to 1 being more smooth.

- Directional order. The direction of the flow is expected to be smooth and coherent for glacier flow. To express this the parameter S is used, defined as:

$$S = \langle \frac{3cos^2(\theta) - 1}{2} \rangle \tag{3.4}$$

  with θ being the angle between a vector and the reference vector (director) running from 0 to 180 degrees. S is a value of 0 to 1. S=0 means there the directions are isotropic and random, while S=1 means all the directions are aligned along the director. The director can be selected from an independent velocity field, other source of flow direction information or set from visual inspection.

These performance indicators provide tools to apply the two tests and select the correct parameters for these tests.

In this research only the SNR threshold is selected based on the performance indicators. The spatial and temporal success rate should be as large as possible, while the MAD on stable areas should be as small as possible. At the same time the dispersion should be as small as possible. The SNR threshold is based on these criteria. The maximum number of combinations used to estimate the velocity at an epoch is chosen as large as possible within the temporal baseline limit of two years. This two year is chosen because the correlation is poor afterwards, introducing more faulty results. The maximum number of used combinations is set at 30 combinations. The minimum amount of displacements to calculate the median is chosen to be 5 or 20 for the SPSM or NNSM respectively. These values are selected, because this is a minimal size of set of values from which a median can be calculated. 15 for the NNSM because the set of values from which to select the median is much larger. The next section describes how the base method and SM methods are tested and compared.

3.2.3. Methods of verification and testings

The selection between the SM, SFSM and NNSM methods will be done based on the success rate and MAD on stable ground as well as visual inspection of the velocity fields. The SNR threshold and reference window size will be chosen based on the success rate, MAD on stable ground and dispersion.

Testing the selected method against the base method is done by looking at all of the performance indicators of both methods, but also requires validation. The first type of validation is the velocity fields from the base method itself. The BM and SM velocity field with the highest success rate and from a cloud free scene are selected. After post-processing both velocity fields, the velocity field from the base method is assumed correct. The two velocity fields are now compared for two test areas in both case study areas. The test areas for the Karakorum region are the Baltoro and the Kyagar glacier, shown in figure 3.9. The Kyagar glacier is a dynamic glacier, thus experiences surges and changes in velocity. The Baltoro glacier is one of the longest glacier outside polar regions. The test areas for the Everest

regions are the Ngozumpa and the Kangshung glacier, these are both debris covered glaciers, see figure 3.10. Quantification of similarity of the the BM and SM field is tricky, because the assumed reference field is not necessarily correct. Therefore the performance test is executed by comparing the speed distributions on glacier between the reference BM velocity field and the tested SM velocity field. Also, the Root Mean Squared Error (RMSE) is computed between the BM and SM results. Furthermore the order parameter S is inspected for the test areas and a visual inspection is applied.

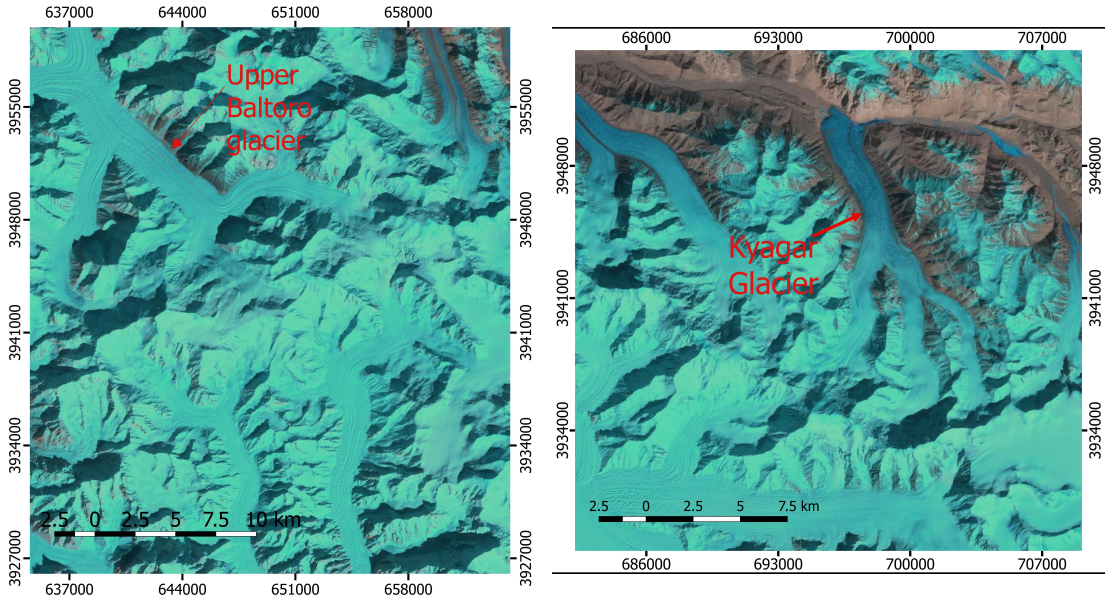


Figure 3.9: The two test areas selected from path/row 148/035, the Karakorum region. The left area is around the Upper Baltoro glacier, the right area around the Kyagar glacier.

The second method of comparison and testing is the external validation. This will be discussed in the next section.

3.2.4. Post-processing

As covered before, the post-processing is not only used for outlier removal, but also as validation. Two types of post-processing can be used: temporal and spatial filtering. The speed of a glacier can be expected to be relatively stable, without glacier surges some small seasonal and temporal variations can be expected, but not large jumps between two epochs with a small temporal baseline. In [35] a temporal Bonferonni outlier test was used, but this type of time series filtering does not work well for dynamic glaciers or sparse time series. This research therefore uses a rolling mean filter of a pixel, the 8 neighboring pixels and the 5 temporally adjacent pixels and their 8 neighbors, which is a maximum total of 45 pixels. This type of filter already restricts the results, as it expects the results to be smooth. The difference between the rolling mean of the pixel and the actual value of the pixel must be smaller than a threshold. This threshold is set at 70 m/yr. If the difference is larger than this threshold, the velocity is masked.

The spatial filtering is more common in this field of research and more methods exists. Most of these are low pass filters or Hampel filters, meaning that they test a pixel based on the neighboring pixels using a Gaussian or square window [17]. Other methods use reference velocity fields to test for outliers [35] or filters based on flow direction. This research will not use a spatial filter. The proposed methods already implements a spatial filter as they require a minimal stack size to calculate the median from.

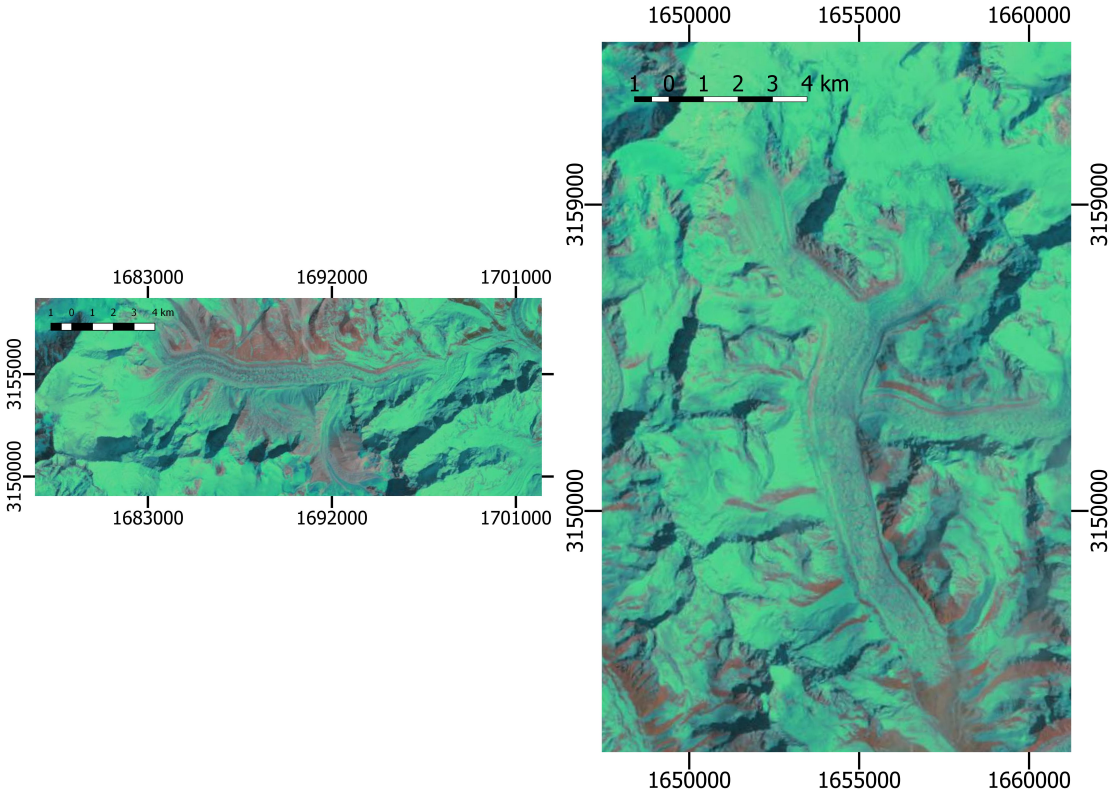


Figure 3.10: The two test areas selected from path/row 140/041, the Everest region. The left area is the area around the Kangshung glacier, the right area around the Ngozumpa glacier.

3.2.5. External validation

The external validation could be provided from four sources: similar optical flow fields (extracted from other satellite optical missions), flow fields extracted from radar satellite missions, GPS velocities or UAV velocities. There are no time series analysis for this region with a high temporal density, thus this validation is for displacements at specific epochs.

Problems with this type of external validations are the different resolutions, source or the different epochs used. The validation can thus only be used as an indication and not quantify performance of the methods.

Results from the same four glaciers as selected in the previous section are compared to results from [31], [32] and [36]. This validation encompasses all three types of sensors and a wide range of methods as well as the complete period of analysis. Using these results as a reference, velocity fields will be compared, but also velocity profiles of glaciers as shown in figure 3.11.

3.2.6. Precision and error budget

There are two kinds of errors in the velocity fields. When a displacement is found it either is the correct displacement (or close to it) or it is a faulty displacement. One type of error is thus a complete faulty displacement. The other error is linked to the precision of the correct results.

The error linked to the precision of a correct result is a combination of errors due to the precision of Imgraft, the co-registration precision and orthorectification errors. These propagate to become the precision of the velocity calculation. Standard error propagation rules combine these errors and scale them according to the calculation of the velocity.That is, the velocity is calculated as:

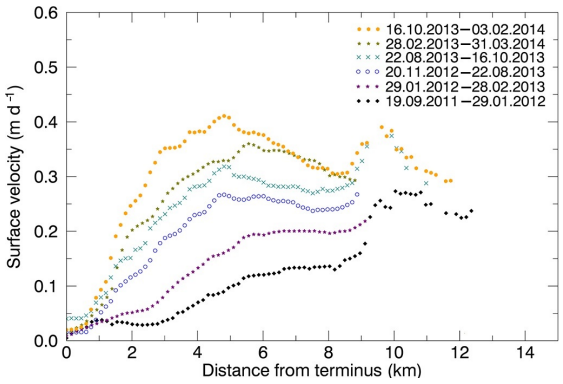


Figure 3.11: Example of velocity profile changes of Kyagar glacier. These results are extracted from TanDEM-X data using offset tracking [36]

$$v = r * (365/T) * \sqrt{(dx)^2 + (dy)^2} \tag{3.5}$$

with dx and dy the displacements in pixels in the x and y-direction respectively, r the resolution of a pixel, T the temporal baselines between the two images to create the velocity field in days and v the speed in m/yr. Assuming the errors of the x and y displacements are uncorrelated, the expression for the standard deviation of the velocity becomes:

$$\sigma_v = 2 * T * \sqrt{\frac{(dx * \sigma_{dx})^2 + (dy * \sigma_{dy})^2}{v}} \tag{3.6}$$

where σ_{dx} is the precision of the displacement in the x direction, σ_{dy} is the precision of the displacement in the y direction and σ_{dv} the precision of the velocity. If the errors in the x and y-direction are assumed similar in magnitude this simplifies to a scaling of the precision of the displacement error with $2 * T$. The precision of the displacement consists of the precision of Imgraft and the co-registration precision. The precision of Imgraft will be calculated and the co-registration error is presented with the Landsat data. Thus the error of the displacement is:

$$\sigma_d = \sqrt{\sigma_{coreg}^2 + \sigma_{imgraft}^2} \tag{3.7}$$

where σ_d is the precision of the displacement, σ_{coreg} the precision of the co-registration and $\sigma_{imgraft}$ the precision of Imgraft. But the precision of the co-registration is the combined precision from the two images used to calculate the displacement, according to:

$$\sigma_{coreg} = \sqrt{\sigma_{master}^2 + \sigma_{slave}^2} \tag{3.8}$$

with σ_{master} the precision of the master image and σ_{slave} the precision of the slave image. When using combinations of displacements to express a displacement (like $d_{1-2} = d_{1-3} - d_{2-3}$) the error also add up in the same way:

$$\sigma_{cd} = \sqrt{\sigma_{df}^2 + \sigma_{d2}^2} \tag{3.9}$$

with σ_{d1} the precision of displacement 1, σ_{d2} the precision of the displacement 2 and σ_{cd} the precision of the combined displacement.

The RMSE of the Landsat scene is calculated by using a similar correlation feature tracking method with respect to a master WRS-2 scene. The RMSE shows no bias or trend, thus here is no relation between the RMSE of two temporally adjacent Landsat scenes [1]. Furthermore for Landsat 7 and Landsat 8 scenes the RMSE is often smaller then 5 meter RRMSE and 55% of the scenes have a RRMSE < 1 meter.

As the feature tracking uses the correlation of a reference window, the error propagation as proposed is not valid. It does give a good indication however what happens with the errors during the velocity calculations.

The proposed SM uses the median to estimate the true value and the MAD to calculate the dispersion, this means that the error propagation is one-on-one. That is, the median selects the middle value of the set when it is odd sized, or the average of the two middle values when the size of the set is even. Thus when the set is odd, the result is a value with an error that is the error as propagated by the velocity calculation. In other words, the estimated value could be a perfect estimate with the same error as the true value. Suppose now the error due to co-registration is 2 meter for a scene pair, a temporal baseline of 16 days would mean that the error due to the co-registration is about 86 meters. When the median would select a velocity results that arises from a combination of displacements, this would even be around 130 meter. As we noticed that the error propagation is not valid, three other error statistics are of interest that indicate the error propagation and the accuracy of the SM.

As discussed, the velocity from the SM method will have a performance statistic, the dispersion, and will be compared to the velocity obtained by the base method, which is assumed correct. This dispersion, the MAD on stable ground and the error between the BM and SM are the three error statistics used in this research.

The dispersion and MAD both express the sample variance according to the central limit theorem. This means that, if no faulty displacements are left in the velocity fields or displacement stacks, the stable MAD and dispersion both express the random errors from the co-registration and Imgraft. The dispersion expresses the random error of a single velocity result and the stable MAD of a complete velocity field. This relation can be seen from figure 3.12, where we can see that a larger MAD on stable ground is related to a larger co-registration error. Thus, the dispersion does not expresses how good the median estimate is compared to the true value as this is dominated by the velocity error, but does describe the error distribution of the estimate and thus of the velocity result.

Figure 3.12: Scatter plot of co-registration error versus MAD on stable ground of all Landsat 8 Everest velocity scenes. The x-axis shows the co-registration error in [m/yr] and the y-axis the stable MAD in [m/yr].

These statistics do not express the performance of the median estimator. This performance is expressed by the RMSE between the velocity of the base method (or other velocity result) and the estimated velocity of the SM. As the proposed testing of the SM includes the comparison of the velocities created by both methods, the RMSE is included in these results as a measure of how well the SM estimates results from the base method.

3.3. Methods summary

There are two parts to this method chapter. The first part describes the core of the setup, it explains how velocity fields are created from image pairs and identifies and sets the parameters of importance for this part.This part of the setup is depicted in figure **??** and consists of:

1. Input: Selection of Landsat multiband image pairs
2. Band selection and PCA for band combinations
3. Application of high pass filter
4. OC feature tracking using the Imgraft feature tracking algorithm
5. Output: x and y-displacement fields and a SNR field.

It also selects two test areas, the Karakorum and Everest in the Himalayas. Finally a method is proposed for dealing with SLC-off Landsat scenes. This method is based on filling the band gaps existing in these scenes with white noise.

The second part describes the proposed method of time series creation and the three variations within this method, and is depicted in figure 3.1. The proposed method works by combining velocity fields to create redundant velocity fields for a temporal baseline, a stack of velocity fields, and is thus called the stacking method. Whereas using a single velocity field with the same temporal baseline is called the base method. The three variations within this method have to do with how this stack of redundant velocities is handled to estimate the true velocity: 1. A median of only same pixel velocities from the stack of velocities 2. A median of only same pixel velocities from the stack of velocities after spatial filtering 3. A median of same pixel velocities and the 8 neighboring pixels from the stack of velocities. The parameters of importance are again identified for this step and it is argued this method could improve performance by creating validation and not filtering out correct values as is a problem with the base method. The performance is how many assumed correct values a method finds on glaciers and how stable these results are, and can be indicated by:

- The success rate
- The dispersion MAD
- MAD of stable ground
- The smoothness of the time series
- The number of values in a time series

Finally an error budget is setup for the algorithm and post-processing methods are selected. An important conclusion from the error budget is the fact that the median estimator has no effect on the precision of velocity results, as the precision does not propagate. Also, the magnitude of the velocity error is dominated by the temporal baseline. The selected parameter values so far are summarized in table 3.3

Region	Sensor	band(s)	high pass filter (pixels)	reference window (pixels)	window spacing (pixels)	maximum velocity (m/yr)
Karakorum	LT5	4,5	1000 meter	12 or 16	6 or 8	500 or 800
Karakorum	LE7	4,5	1000 meter	12 or 16	6 or 8	500 or 800
Karakorum	LC8	8	1000 meter	12 or 16	6 or 8	500 or 800
Everest	LT5	4,5	1000 meter	12 or 16	6 or 8	500 or 800
Everest	LE7	4,5	1000 meter	12 or 16	6 or 8	500 or 800
Everest	LC8	8	1000 meter	12 or 16	6 or 8	500 or 800

Region	Submethods	Nr combinations	Max temporal baseline [#]	Minimal stack size	SNR threshold [-]	Temporal filter threshold $[\Delta m/yr]$
All	SFSM & SM	30	730 days	5	To be determined	70
All	NNSM	30	730 days	15	To be determined	70

Table 3.3: Algorithm parameter settings for test areas and sub-methods. 12 or 16 means that these results will both be tested and compared in the results.

4

Sensitivity analysis & algorithm selection

This chapter is divided in four sections. First of all results from the Imgraft precision testing and SLC-off white noise solution are presented. The second part is parameter tuning of the SNR threshold and reference and searchwindow sizes. The third section presents results from the SM sub-methods comparison and the selection of the NNSM. Finally issues concerning the functioning of the NNSM are presented and discussed.

4.1. Imgraft Precision & SLC-off scenes

4.1.1. Imgraft precision

To complete the error budget and take a look at possible other errors in the Imgraft algorithm, the Imgraft precision is calculated.

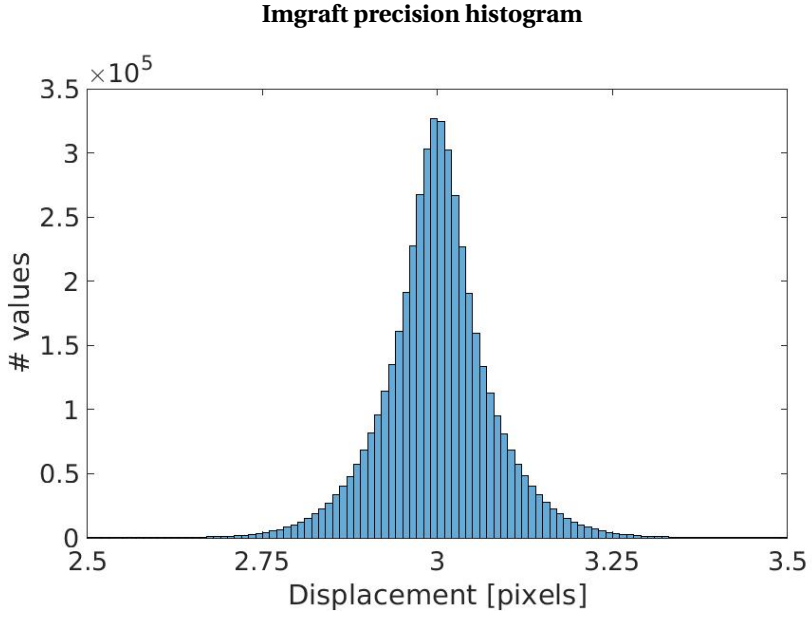


Figure 4.1: Histogram of displacement in x-direction between two equal, cloud free Landsat 8 scenes. One of these scenes is displaced 3 pixels. The results with perfect precision should thus be 3 pixels.

In figure 4.1 a histogram of the Imgraft precision is plotted. This is done by selecting a cloud free Landsat 8 scene with 15 meter resolution and copying this. The copy is displaced by 3 pixels in the positive x direction. The Imgraft algorithm parameters are set to the selected parameters. The result should

give a displacement of 3 pixels if the precision of Imgraft were to be perfect.

The statistics of this distribution thus give an indication of the precision and accuracy of Imgraft. The mean of the distribution is 3 pixels. The σ is **0.0886 pixels**. This is about 1/10th of a pixel. The algorithm also returns a few large outliers with a SNR between 0-2. All these outliers reside in straight lines along the edges of the results. It should be noted that these outliers on the edges of the images might affect the MAD on stable ground and affect the selection of the SNR threshold.

The lower limit precision of 1/10th of a pixel is chosen for all pixels.Thus, an Imgraft error of 0.1 pixel can give an error of 60 m/yr meters for a 16 days temporal baseline using a resolution of 15 meter.

4.2. Landsat 7 SLC-off results

The Landsat 7 SLC-off scenes are processed in two ways: first by application of orientation correlation without white noise gap filling and second by applying a white noise gap mask. It is tested by comparing single velocity pair results and their performance. Because the used correlation method is OC, the correlation does return results for SLC-off scenes. The results for a single velocity field are compared in figures 4.2 and 4.2.

a. White noise filled **b. Data gap masked**

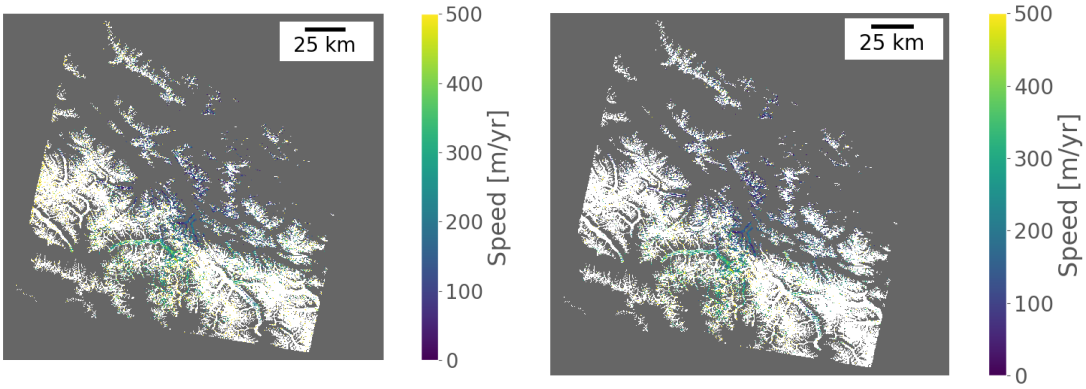


Figure 4.2: Velocity fields from Landsat 7 SLC-off images in winter 2004 with a 32 days temporal baseline. The images are filtered with a SNR threshold of 3 and are processed using the same algorithm parameters, with a maximum velocity of 500 m/yr. The grey areas indicate the non-glacier masked areas. (a) white noise gap filled velocity field. (b) . Non-gap filled velocity field. Enlarged figures can be found in Appendix A A.1 and A.2.

The figures in 4.2 show both SLC-off velocity results for the complete WRS-2 Karakorum scene. All non glacier areas are masked and indicated by the gray area. All the white areas are the glacier masks in the Karakoram region.

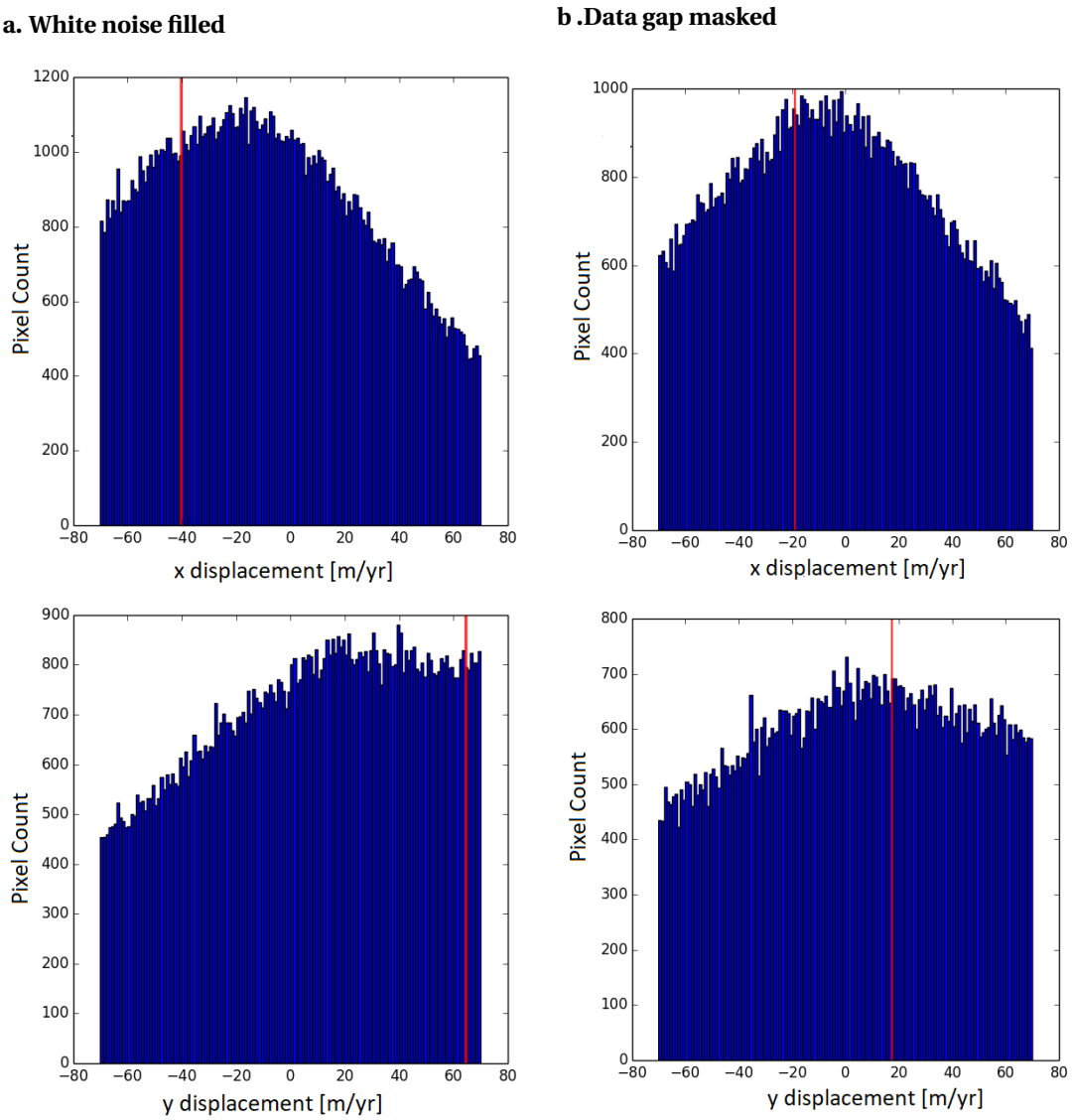


Figure 4.3: Displacement distributions in x and y direction on stable ground of Landsat 7 SLC-off white noise filled scene of (path/row 148/035). (a) shows the distribution of the white noise filled product. The median of the displacement in the x direction is -41.3 m/yr, the median in the y-direction is 62.6 m/yr. (b) shows the distributions of non-gap filled products. The median of the displacement in the x direction is -20.4 m/yr, the median in the y-direction is 18.8 m/yr

Visual comparison of the two BM velocity fields in figure 4.2 provides little information on the performance difference. In the middle of the scene, where there are no gaps or small gaps in the data, assumed correct velocities are found on the glaciers. Towards the edges of the scenes, speckle is visible in both images and almost no velocities are correct. This is because the gaps in the North- South direction are wider than the size of the search window. The performance of both techniques is addressed in the figures of 4.3. This figure show the distribution of the x and y-displacement on stable ground for both the white noise and gap mask method. The displacement in the x direction shows a hint of normal distribution, as expected, for both gap masking techniques. The displacement in the y direction however is very biased for the white noise technique. Overall the displacements from the white noise technique are more biased, especially in the y direction, and have a larger stable MAD. **Based on these single pair performance results, the SLC-off scene data gaps are not white noise filled.**

It is clear that the results from SLC-off scenes show little results around the edges, where the gaps are wide and have more assumed correct results when the gaps get smaller. The correct displacement

in the y direction can only be acquired if the displacement is smaller than the width of a data area or the displacement crosses a data gap. Furthermore the search window must be large enough to cover this gap. Also, the displacement in the y-directions performs very poorly when white noise filling is applied. One of the main reasons why the white noise filling as applied in this research does not function is the limited 8-bit quantization of the data. For the white noise filling to work, the white noise must be very random. The 8 bit quantization however limits the white noise randomness to values from 0-255. This is not random enough and the correlation finds displacements with a high SNR within the white noise filled areas.

4.3. Parameter tuning

In this section the final parameters are tuned. The first subsection selects the window sizes and maximum velocity. The second subsection present results from a sensitivity analysis of the SNR threshold. Both are based on the MAD on stable ground and success rate performance indicators.

4.3.1. Imgraft parameters

One important parameter is the choice of the size of the reference window and linked to this the spacing parameter and the maximum velocity. In this section a window size of 12 and 16 pixels with respective maximum velocities of 500 and 800 m/yr are compared. The performance results are presented in figures 4.4 and 4.5

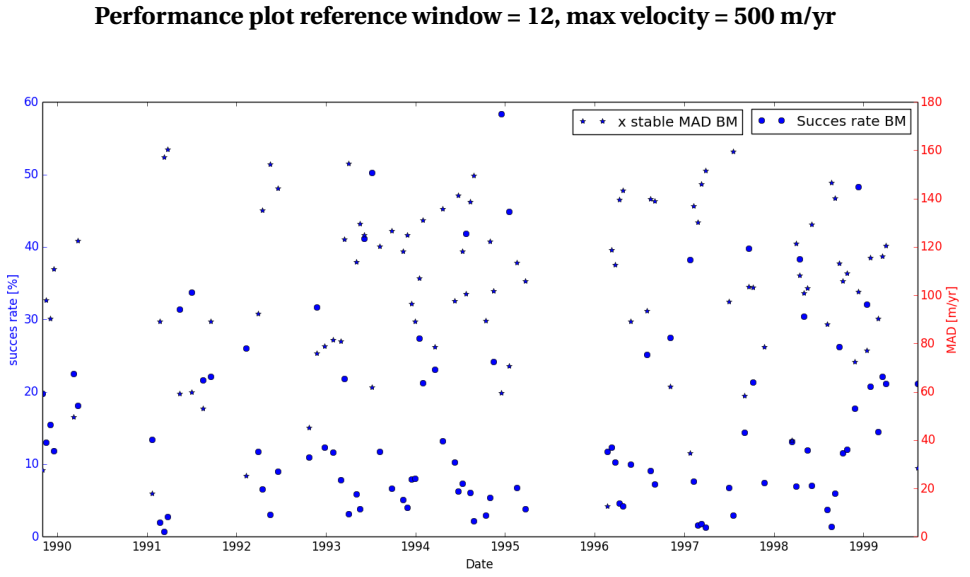


Figure 4.4: Performance plot for time series of Landsat 5 for the Karakorum region (path/row 148/035). The left axis is the success rate percentage, the right axis the MAD of stable area in m/yr. The epoch of a velocity field is the epoch of the slave base pair image used for the velocity field. The velocity fields are calculated with a reference window size of 12 pixels, spacing of 6 pixels and maximum velocity of 400 m/yr and high pass filter distance of 1000 meter.

To compare the results of these two parameter settings the average of the MAD and success rate and the variance of the MAD are calculated and compared in table 4.1.

Reference window size	Mean succes rate	Mean MAD [m/yr]	Variance MAD [m/yr]
12	15.6	102	34
16	16.5	154	122

Table 4.1: Performance comparison of Imgraft parameter settings tested on Landsat 5 TM scenes of path/row 148/035.

The difference between these two methods is also displayed as velocity fields in figure 4.6a and

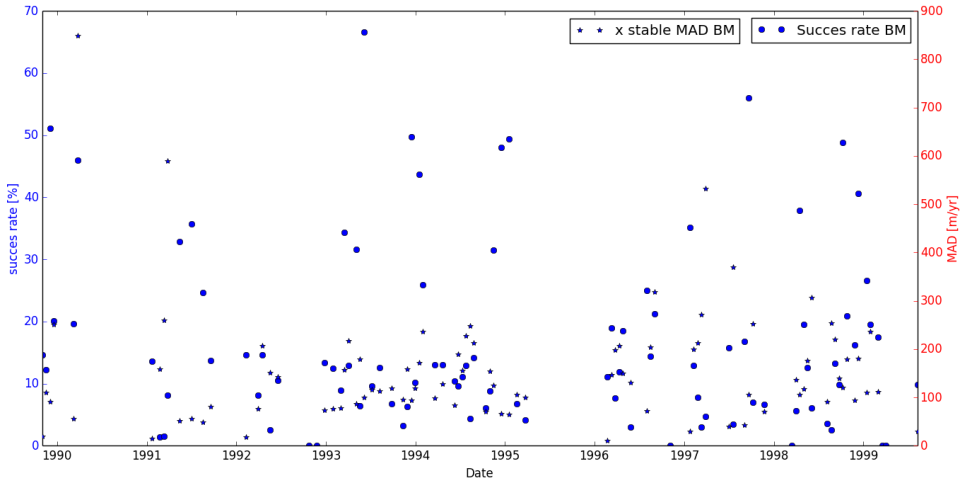


Figure 4.5: Performance plot for time series of Landsat 5 for the Karakorum region (path/row 148/035). The left axis is the success rate percentage, the right axis the MAD of stable area in m/yr. The epoch of a velocity field is the epoch of the slave base pair image used for the velocity field. The velocity fields are calculated with a reference window size of 16 pixels, spacing of 8 pixels and maximum velocity of 800 m/yr and high pass filter distance of 1000 meter.

figure 4.6b. This displays the two velocity fields of the two time series with the highest success rate for the Kyagar glacier and surroundings in the Karakorum region.

The visual inspection and performance show that a reference window of 12 pixels and a maximum velocity of 500 m/yr is more stable while maintaining a similar success rate.The difference in success rate is mostly due to faulty results on semi-stable ground. The larger reference window and larger maximum velocity, e.g larger search window size, thus increase the success rate, but most of this increase in success rate is due to faulty velocities.

The selected reference window size is thus 12 pixels and the selected maximum velocity is 500 m/yr. This selection of a smaller maximum velocity is problematic, because peak velocities from surging glaciers cannot be registered. The use of static parameters for large scale research is problematic anyway as different regions do benefit from different parameter settings, but computationally more efficient.

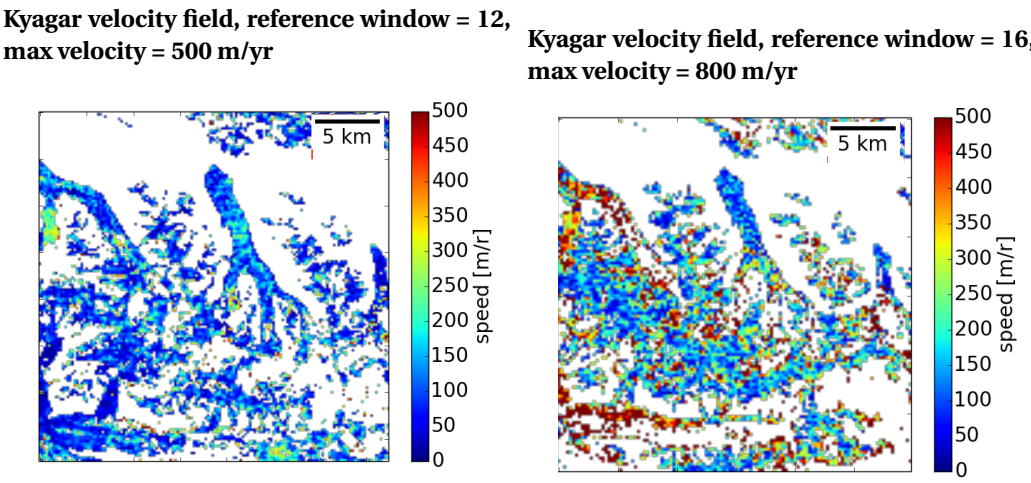
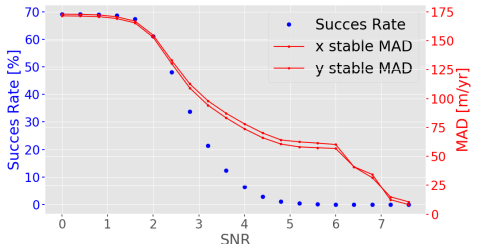


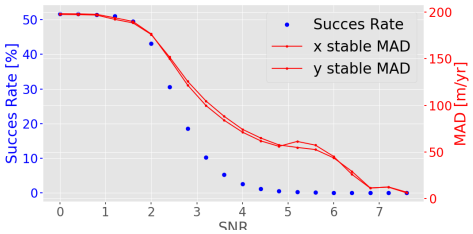
Figure 4.6: Landsat 5 TM velocity fields of the Kyagar domain. (a) shows the velocity field using a reference window size of 12 pixels and maximum velocity of 500 m/yr. (b) shows the velocity field fusing a reference window size of 16 pixels and maximum velocity of 800 m/yr.

4.3.2. Selection of SNR threshold

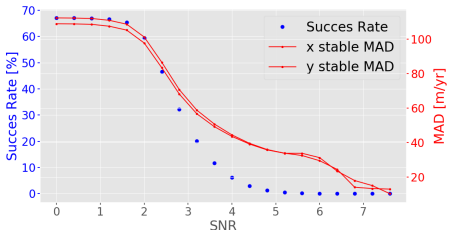
The SNR threshold is the final parameter that is tuned. The average performance results from the base method for all velocity fields per sensor type and path/row are depicted in figure **??**. These plots show the average stable MAD in the x and y direction and the success rate of all time series velocity fields against the SNR threshold. Figure 4.8 shows the same plots, but include the average dispersion.



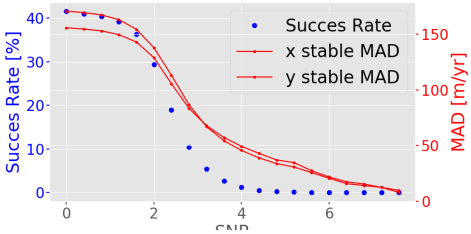
(a) Performance Landsat 5 TM path 140 row 041



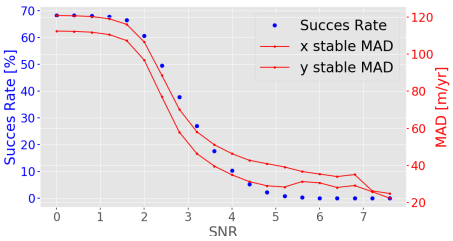
(b) Performance Landsat 5 TM path 148 row 035



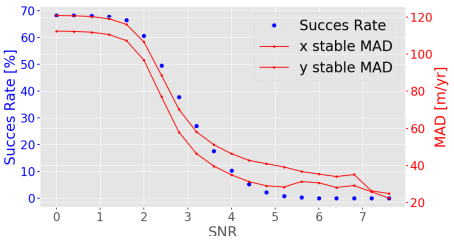
(c) Performance Landsat 7 ETM+ path 140 row 041



(d) Performance Landsat 7 ETM+ path 148 row 035



(e) Performance Landsat 8 OLI path 140 row 041

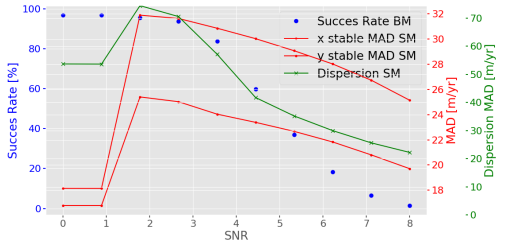


(f) Performance Landsat 8 OLI path 148 row 035

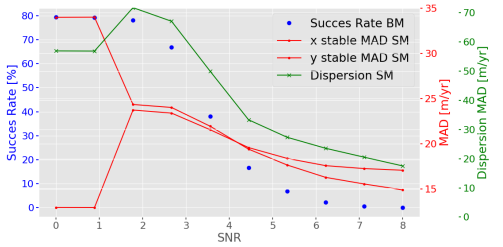
Figure 4.7: Average of success rate and stable MAD as a function of the SNR thresholdfor all types of sensors for both testing regions using the base method.

The average performance results of the NNSM method of 40 velocity fields per sensor type and path/row are depicted in figure 4.8.

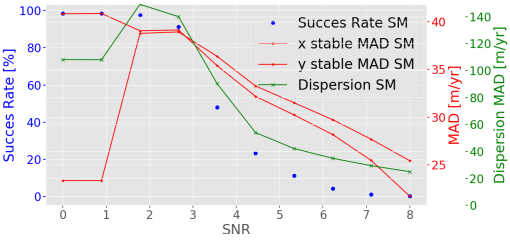
There are several things to note about these results. First of all, the NNSM results are only an average of the first 40 consecutive time series velocity fields and only have 10 SNR threshold data points opposed the the base method method where all velocity fields are used and 20 SNR threshold data points are used. This is because of the large computational effort related to this average performance. Another important note is about the behaviour of the average performance at SNR thresholds below 2. This behaviour can be explained as arising from edge effects. Some results of velocity fields systematically contain large strips around the edges of the scenes where the SNR is not zero and the displacements are very low or systematically high. Sometimes this only occurs along an edge in one direction as in figure4.8b. Also, the results from the NN-stacking method have a very high success rate at very low SNR thresholds. This is because the stacks from which the medians are calculated contain enough values to get a velocity for every pixel, but these velocities are not correct.



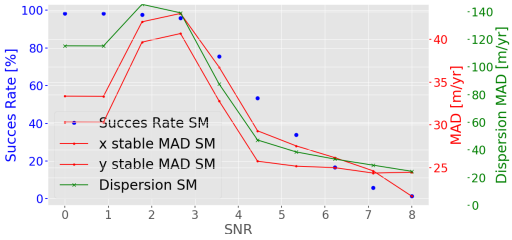
(a) Performance Landsat 5 TM path 140 row 041



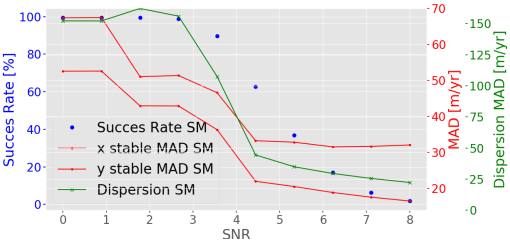
(b) Performance Landsat 5 TM path 148 row 035



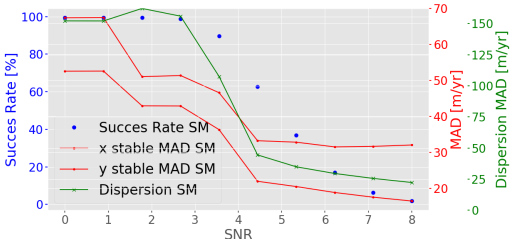
(c) Performance Landsat 7 ETM+ path 148 row 035



(d) Performance Landsat 7 ETM+ path 140 row 041



(e) Performance Landsat 8 OLI path 140 row 041



(f) Performance Landsat 8 OLI path 148 row 035

Figure 4.8: Average of success rate and stable MAD as a function of the SNR thresholdfor all types of sensors for both testing regions using the NN-stacking method.

The selected SNR threshold must be applicable to all scenes in the Himalaya for a specific sensor type and must be a realistic value for comparing the results from the base method and stacking method. The ideal selection of the SNR threshold would thus be the same threshold for both methods, but this would be an unfair comparison as we can already see the difference in performance from this sensitivity analysis. Thus, two different SNR thresholds are selected based solely on the sensitivity plots of the respective method.

As this average performance does not include a temporal filter, the threshold is chosen lower than advised from the performance. The temporal filter is then expected to filter out any faulty velocity results that are left. The dispersion and MAD should be as low as possible, while the success rate should be as high as possible. This results in a SNR threshold of 4 for all three sensor types for the NN-stacking method. The selection of the SNR threshold for the NN-stacking method is more clear as it also includes the dispersion indicator. The SNR threshold for the base method is harder to select as there is not really a point at which the difference between the success rate and stable MAD is optimal. A SNR threshold of 3 is selected for the Landsat 5 and 7 scenes, while a SNR of 2.5 is chosen for the Landsat 8 scenes. This is summarized in table 4.2.

Method	Landsat type	SNR threshold
NNSM	Landsat 4&5	4
NNSM	Landsat 7	4
NNSM	Landsat 8	4
BM	Landsat 4&5	3
BM	Landsat 7	3
BM	Landsat 8	2.5

Table 4.2: SNR threshold selection per method and Landsat type.

4.4. Method selection

As stated in the section "Proposed method", three stacking method sub methods have been tested and compared in this research. This section will compare these methods based on visual inspection and the quality indicators and present why the NN-stacking method is the best sub method.To recap, the three methods are:

- Median of all velocities of same pixel locations in stack of velocity fields (SPSM).
- Median of all velocities of same pixel locations in stack of velocity fields after application of spatial filtering to filter out noise (SFSM).
- Median of all velocity fields in stacks of same pixel locations and nearest neighbors (NNSM).

4.4.1. Comparison single pixel stacking method and nearest neighbor stacking method

The comparison is based on both a visual and performance plot of the methods. In figure 4.9 and figure 4.9 and table 4.3 the performance indicators success rate and stable MAD are compared for every epoch for the Landsat 5 TM scenes of path/row 148/035. The epoch is expressed as the date of the slave image of the velocity base pair.

Performance statistics NNSM and SPSM

Method	Mean success rate [%]	Mean stable MAD [m/yr]	Variance stable MAD [m/yr]
SFSM	21.6	107	49
NNSM	25.4	77	40

Table 4.3: Performance results of single pixel stacking method and NN-stacking method for Landsat 5 velocity time series from the Karakorum region (path/row 148/035). The MAD and variance of the MAD is in the x direction only.

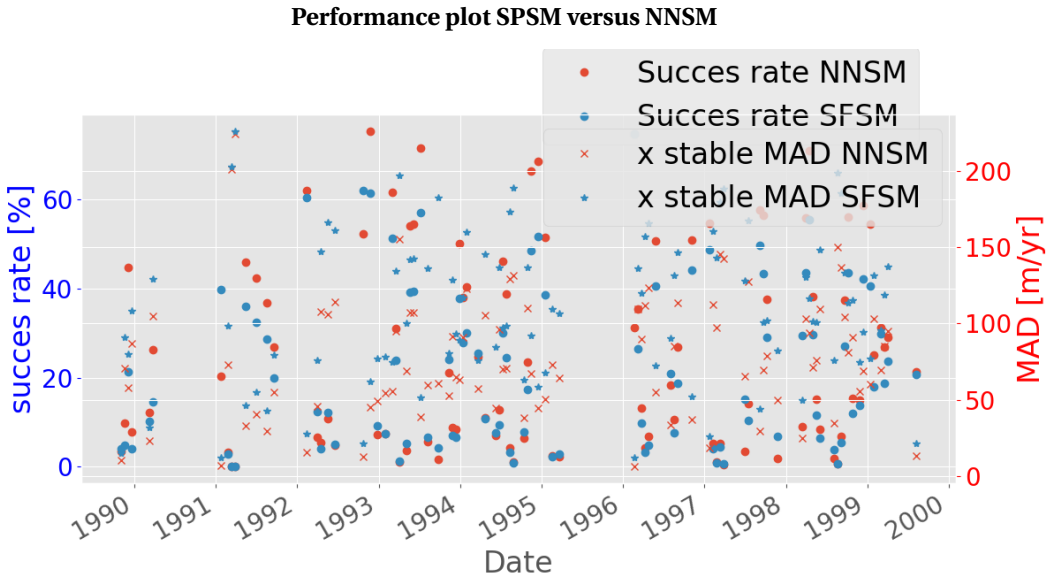


Figure 4.9: Performance plot comparing the SPSM and NNSM Landsat 5 time series of the Karakorum region (path/row 148/035). The left axis is the success rate percentage, the right axis the MAD of stable area in m/yr. On the x-axis the date of the slave image of the base image pair.

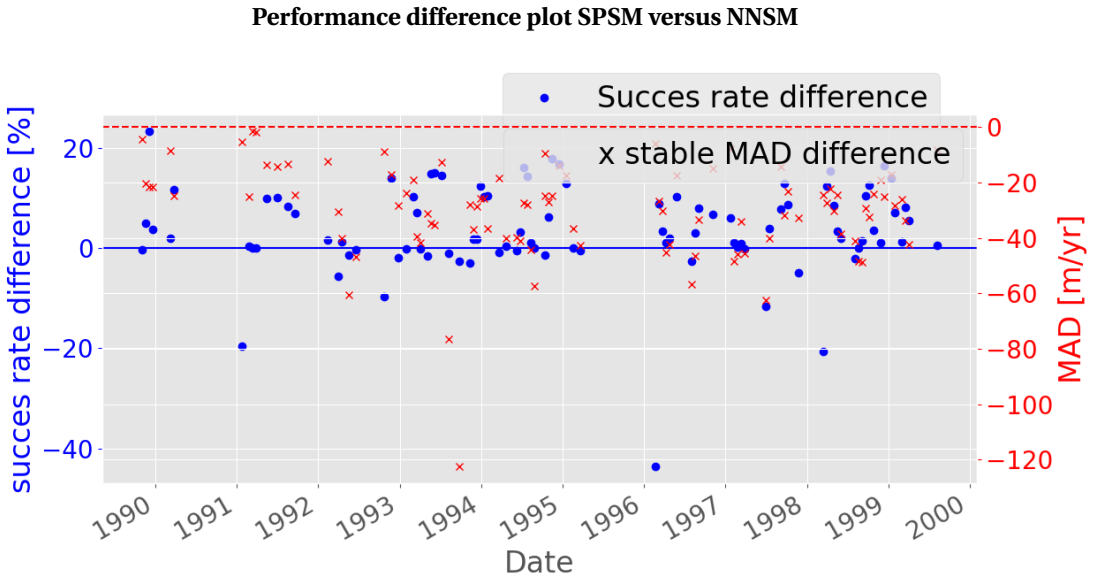
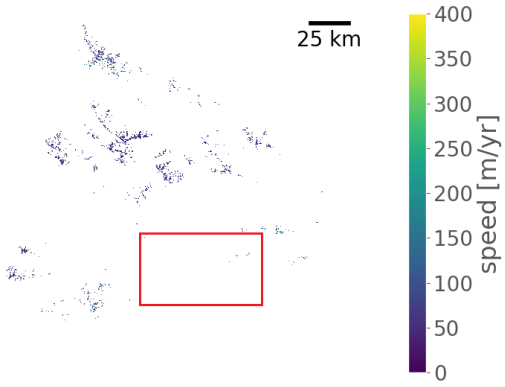


Figure 4.10: Performance difference plot comparing the SPSM and NNSM, e.g NNSM - SPSM, Landsat 5 time series of the Karakorum region (path/row 148/035). The left axis is the success rate percentage, the right axis the MAD of stable area in m/yr. On the x-axis the date of the slave image of the base image pair.

Figure 4.9 shows the absolute performance of the time series from both methods. Figure 4.10 shows the difference in performance, e.g NNMS - SPSM. The positive success rate difference thus means that the success rate from the NN-stacking method higher. A negative stable MAD difference shows that the the stable MAD is lower for the NNSM. The performance of the NNSM is better according to the statistics and plots. From figure 4.10 we can see that the difference between success rates is mostly positive, meaning that the success rate of the NNSM is larger. The MAD difference is negative for most epochs, thus the MAD of the SPSM is larger.

a. Karakoram velocity field from NNSM



b. Karakoram velocity field from SPSM

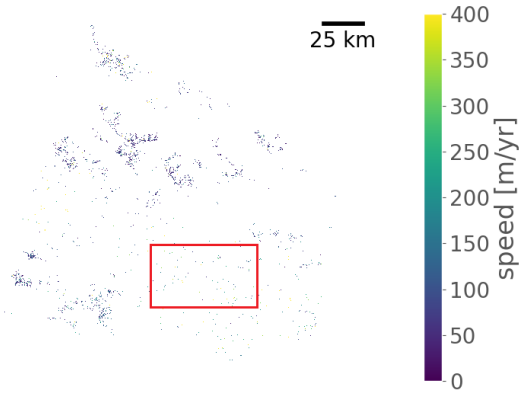


Figure 4.11: Landsat 5 velocity field of Karakoram region (path/row 148/035) created by the (a) NNSM (b) SPSM.

In figure 4.11 the velocity results of the complete scene of path/row 148/035 from 1992/11/29-1992/12/28 are presented for both methods. One of the original scenes is cloud free, the other original image is shown in figure 4.12.

Original Landsat scene 1992/12/28

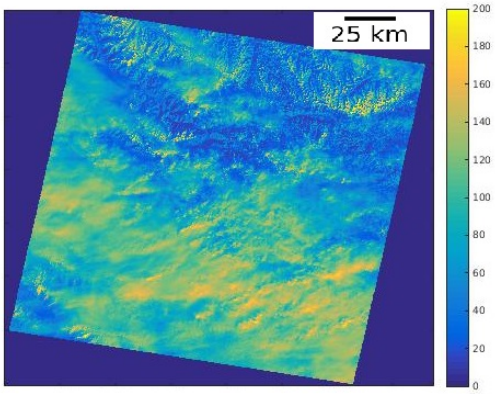


Figure 4.12: Original Landsat 5 scene of path/row 148/035 of 28 December 1992. Clouds show in the bottom area of the scene.

Inspection of these velocity fields provides more information on both performances. The area indicated in figure 4.11 corresponds with clouds in one of the original Landsat images, see figure 4.12. Thus no velocities are preferred and expected in the cloudy area. The NNSM filters most of the velocities in this area. Results from the SPSM however show quite some speckle in this area. This is unwanted and combined with the better performance of the NNSM, **the NN-stacking method is selected as a better method**. The SFSM might prevent this speckle as a spatial filter is applied. Results of this method are presented in the next section.

4.4.2. Comparison spatial filter stacking method and nearest neighbor stacking method

The same comparison is applied to the SFSM and NNSM methods.Table 4.4 summarizes the performance indicators of the two methods. The performance plots can be found in Appendix A, figure A.3 and A.4.

Method	Mean succes rate [%]	Mean stable MAD [m/yr]	Variance stable MAD [m/yr]
SFSM	19.9	105	47
NNSM	25.4	77	40

Table 4.4: Performance results of spatial filter stacking method and NN-stacking method for Landsat 5 velocity time series from the Karakorum region (path/row 148/035). The MAD and variance of the MAD is in the x direction only.

Again the performance of the NN-stacking method is better. The success rate of the spatial filter stacking method is lower than the single pixel stacking method, this is because some of the noise indicated in figure 4.11b is filtered out by the spatial filter. Inspection of the performance shows that the performance of the NNSM however is still better than the SFSM performance. Furthermore the application of a spatial filter next to the SNR filter is computationally very expensive.

Based on results from the previous and this section the NN-stacking method is selected as the best method. The behaviour and issues of the complete algorithm and in particular the NNSM are presented and discussed in the next section.

4.5. Functioning and issues of NN-stacking method

This section shows and discusses the functioning and issues of the NNSM. In this section the effects of shadows, clouds and snow cover on the velocity fields are discussed. It also presents issues of the NNSM and the complete algorithm.

The most important thing to note is the smoothing effect of the NNSM. As the NNSM uses neighboring pixels to estimate the velocity, the velocity field is smoothed. This is an effect that is desirable too some degree. There are however two things that need to be taken into consideration. The first is over-smoothing: while smooth behaviour in velocity fields is somewhat desired, there is a danger of over-smoothing the results and thus obtaining biased velocity fields or removing small scale velocity features. The second is the smoothing in combination with the temporal filter. The temporal filter also requires the velocity field to be smooth. An artificially smooth velocity time series could thus be created by choosing a small threshold for the temporal filter and a SNR that is too low. An extreme example of this is the velocity field presented in figure 4.13 where the SNR threshold is set at 1. This Landsat 5 TM velocity field of the domain around the Kyagar glacier shows a velocity field of close to zero without spatial variations.

Another important result from the NN-stacking method is the handling of shadows, clouds and snow cover. In figure 4.14b two areas are indicated corresponding with shadows and snow cover. The correlation is affected by shadows and snow cover and often returns faulty velocities for these areas. From figure 4.14a we see that these areas are masked when using the NNSM, this is desired from this method and an indication that the algorithm functions well.

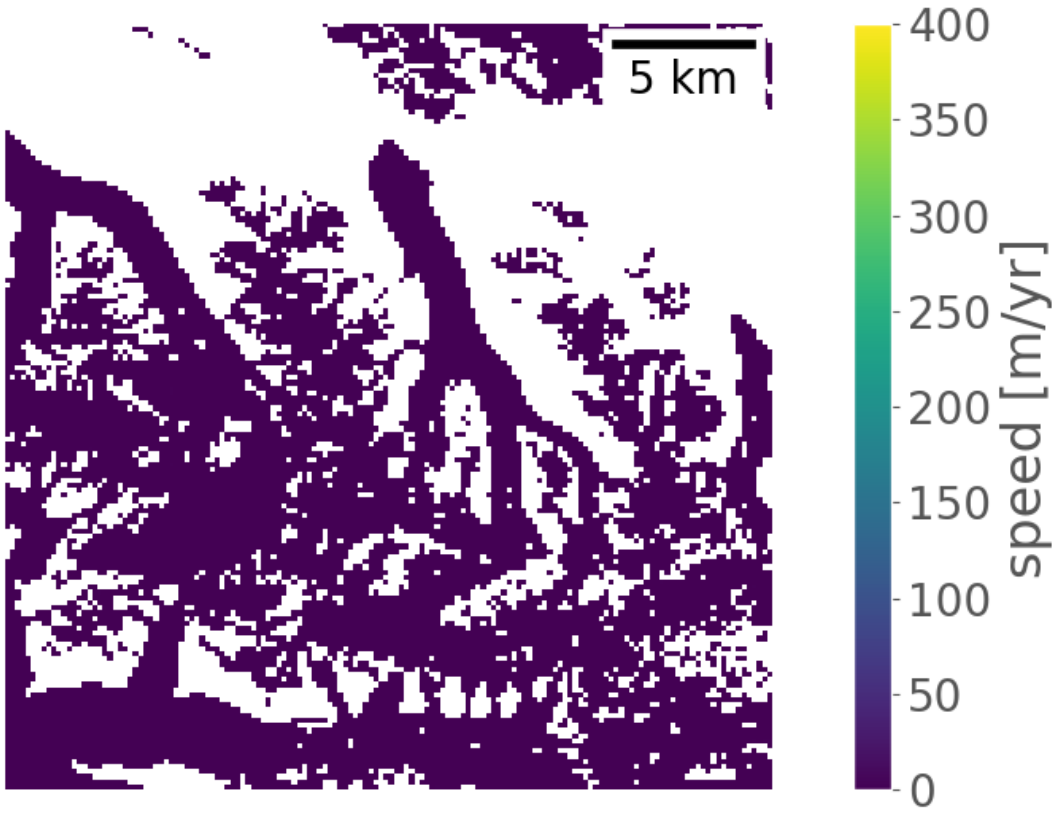


Figure 4.13: Landsat 5 velocity field of the Kyagar glacier domain obtained from the NN-stacking method filtered with a SNR threshold of 1.

a. Landsat 5 Baltoro velocity field

b. Landat 5 Baltoro image from 20 December 1996.

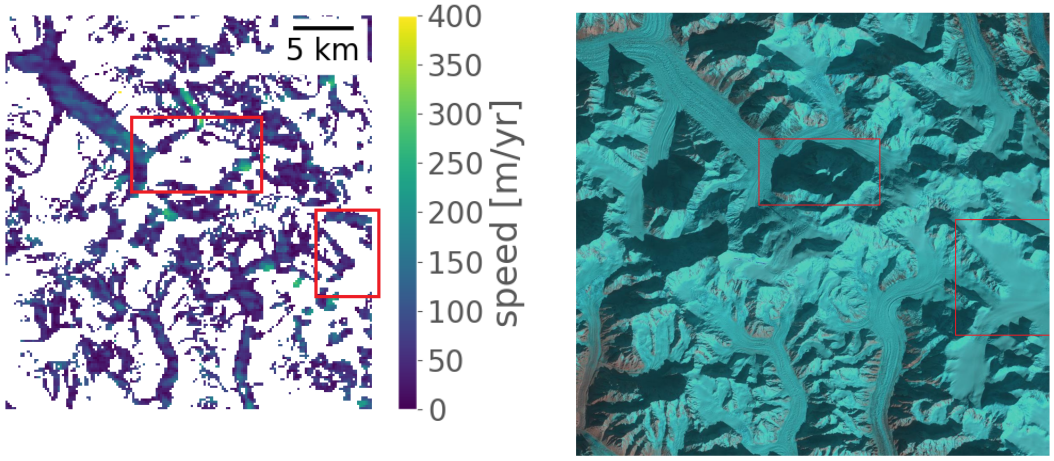


Figure 4.14: Comparison of Baltoro velocity field with original Landsat scene domain. (a) shows the Landsat 5 velocity field from the NN-stacking method (b) shows the Baltoro domain from the original Landsat scene from 20 December 1996.

4.5.1. Velocity field results

The results from the sensitivity analysis show one more issue not related to the NNSM, but the core method (velocity fields) itself. When comparing the base method sensitivity plots to the result from [9] as shown in figure 4.15 some important differences are noticed. The stable MAD of the velocity fields in this regions is much larger. This can be explained by the short temporal baselines used in this research which magnify errors thus enlarging the stable MAD. The other difference is the general effect of the SNR on the results. At a SNR threshold of 6 the success rate is still 20 % in figure 4.15, while the success rate has already dropped to zero for a SNR threshold of 4.5 in results from figure **??**.

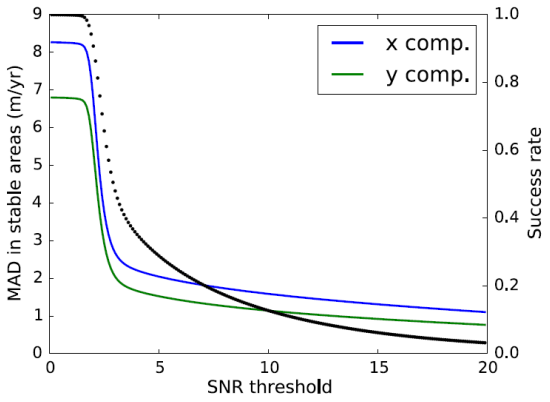


Figure 4.15: Sensitivity analysis from [9] plotting the success rate and stable MAD as a function of the SNR threshold.

There are two options why this could happen. First of all, the Imgraft parameter settings might be the cause. However, the comparison of parameter settings in section "Imgraft parameters" indicates that the effect of a change of Imgraft parameters is very small for the success rate. The second option would be the functioning of Imgraft itself. Imgraft was selected without testing and the selection was based on the precision of the algorithm. The results from this research show however that the behaviour of Imgraft regarding filtering and the signal-to-noise ratio is strange and should be tested.

4.6. Summary

This chapter should answer the sub question **Which combination of feature tracking techniques, parameters and algorithm parameters gives the best result in the areas of research?** but also studies the functioning of the selected method.

The precision of Imgraft was quantified at 0.1 pixel. This also validates the functioning of Imgraft. Based on the results from comparing the Landsat SLC-off gap solutions it is chosen to not fill the data gaps with white noise.
A reference window of 12 pixels and maximum velocity of 500 m/ yr is selected based on the performance of the base method Landsat 5 time series. A higher maximum velocity allows for better monitoring of dynamic and surging glaciers, but the performance decreases. The selection of the SNR threshold is based on the average performance of both the base and NN-stacking method and is summarized in table 4.2. The selection of the SNR threshold is easier based on the NN-stacking method results, because of the added dispersion indicator and bigger difference between performance indicators.

The NNSM is selected as the best sub-method. The performance is better, but it is more important that this method is better in filtering outliers. The NN-stacking method smooths results because it also incorporates spatial information into the velocity estimate. When a correct SNR threshold is selected however the algorithm does recognize shadows, clouds and snow cover and is thus suitable for large area automated analysis. Finally it was argued that the Imgraft algorithm might have a flaw regarding the signal-to-noise ratio. appe

5

Results: Karakoram

In total 5734 velocity fields have been produced for this region. 1500 Landsat 8 velocity scenes, 2928 Landsat 7 velocity scenes and 1306 Landsat 5 velocity fields. The total processing time was around 10 days, of which 7 days for Landsat 8. These velocity fields are combined using the proposed method to create a consecutive dense time series of 96 velocity fields for Landsat 5, 186 velocity fields for Landsat 7 and 86 velocity fields for Landsat 8. The Karakoram region has many glaciers, with a large range of velocities and a highly varying topography. Furthermore many of these glaciers are dynamic, meaning that they often see changes in velocities or undergo a glacial surge (sudden increase of speed). The two studied glaciers, Baltoro and Kyagar have velocities ranging from 0-300 m/yr, but maximum velocities have been measured of > 700 m/yr.

This chapter first presents validation of the NN-stacking method in this region, both by testing against the base method and by external validation. Then it presents performance results of both methods and finally shows time series and time series information.

5.1. Validation of NN-stacking method

5.1.1. Base method validation

The validation of the base method was based on assuming the results from the base method are correct. The results compare velocity fields of the Kyagar glacier for Landsat 5 in figure 5.1 and the velocity fields of Landsat 8 for the Upper Baltoro region in figure 5.2.

The velocity field of the Kyagar glacier from 1992/10/24 to 1992/11/25, as indicated by the red box in figure 5.1a, shows East-West lines of velocity differences. This can be seen in all velocity fields created from Landsat 5 scenes in the Karakoram region. The feature tracking displacement result in the x-direction of the same area, as shown in figure 5.1b, shows the same lines. The direction of these lines coincide with the scan line direction of the TM scanner. It is therefor likely that these lines are due to scan line artifacts. The magnitude of the error introduced by these artifacts is quite large (~0.7 pixel) and the small temporal baseline magnifies this error. It is also not know if these lines shift between scenes, and therefor all Landsat 5 TM results for this region are deemed unfit for time series analysis. The Landsat 8 velocity field of the Upper Baltoro domain from 2016/02/13-2016/02/29, figure 5.2, shows similar results for the NN-stacking method and base method. The Upper Baltoro glacier, indicated by the red box, shows similar velocity and velocity distribution for both methods. The velocity field of the NN-stacking method however, filled the gaps on the glacier that are visible from the base method velocity. This is also true for areas surrounding the glacier, which are indicated as glacier by the glacier mask, but are actually icy/snowy mountain tops or slopes with a year round snow cover, indicated by the blue box in figure 5.2b. The NN-stacking method fills either completes (fills) these areas or removes any points left in them.

The accuracy of the NN-stacking method is expressed as the pixel-wise RMSE between the NN-stacking method velocity and base method velocity in table 5.1 and the speed distribution is depicted

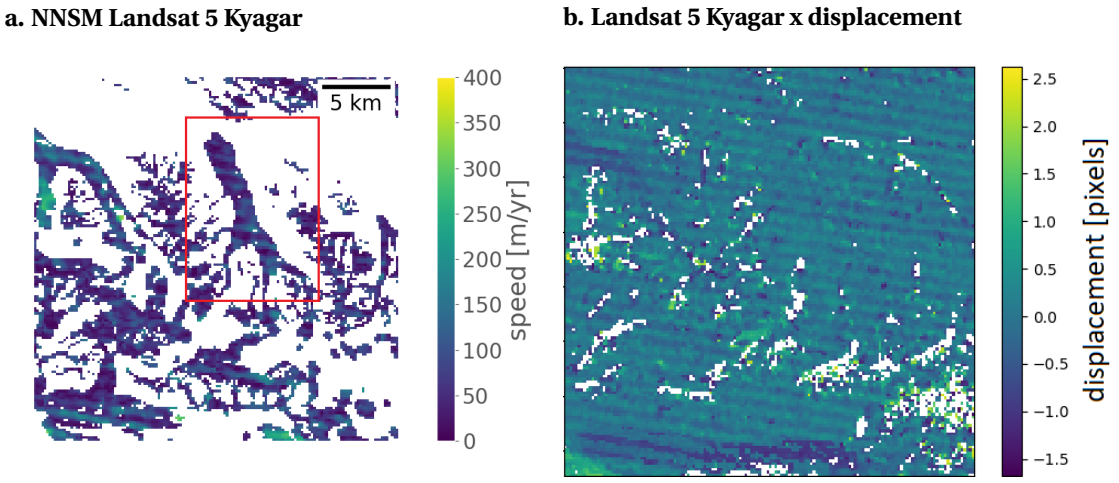


Figure 5.1: (a). Velocity field of domain around the Kyagar glacier from 1992/10/24 to 1992/11/25 for the NN-stacking method. (b) shows the displacement field in the x direction of the same domain.

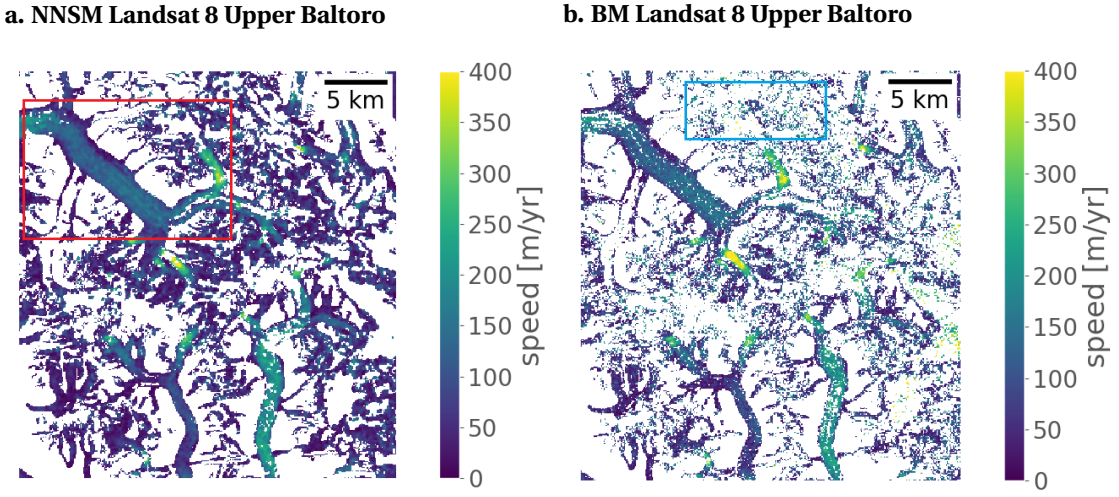


Figure 5.2: Landsat 8 velocity fields of domain around the Upper Baltoro glacier from 2016/02/13-2016/02/29 for both the NN-stacking method and base method. (a) shows the NN-stacking method results, (b) shows the base method results.

in figure 5.3 for the Landsat 8 Baltoro domain. The RMSE is expressed in m/yr and includes all the glacier masked areas, thus also the icy slopes around the glacier which might offset the RMSE results.

The RMSE is the lowest for Landsat 8 results and in the order of 35-45 m/yr. The difference is evenly distributed over the glacier mask, but the velocity is almost always lower (biased) for NNSM results. That is, the RMSE consists of $bias(y_{NNSM}, y_{BM})^2 + var(y_{NNSM}, y_{BM})$ and for a random error only the variance term would add to the RMSE. From figure 5.3 and the velocity difference we can see however that the velocity estimate from the NN-stacking method is biased towards lower velocities. As the error is evenly distributed, the RMSE is also 35-45 m/yr on slow moving glaciers (< 50 m/yr) which is large compared to the velocity of the glacier. Finally figure 5.3 also shows that the NN-stacking method obtains more results in the Baltoro region, but from visual inspection we see that most of these results are again snowy/icy areas around the flowing glaciers.

It was assumed that the results of the base method are good and the results of the NN-stacking method can be validated from the base method. Considering the large RMSE between methods however, external validation provides better validation options.

Landsat	RMSE Baltoro [m/yr]	RMSE Kyagar [m/yr]
LT5	49.9	43.6
LE7	59.6	42.4
LC8	45.6	36.9

Table 5.1: Root mean square error between the NN-stacking method velocities and the base method velocities for all Landsat types and the Baltoro and Kyagar glacier.

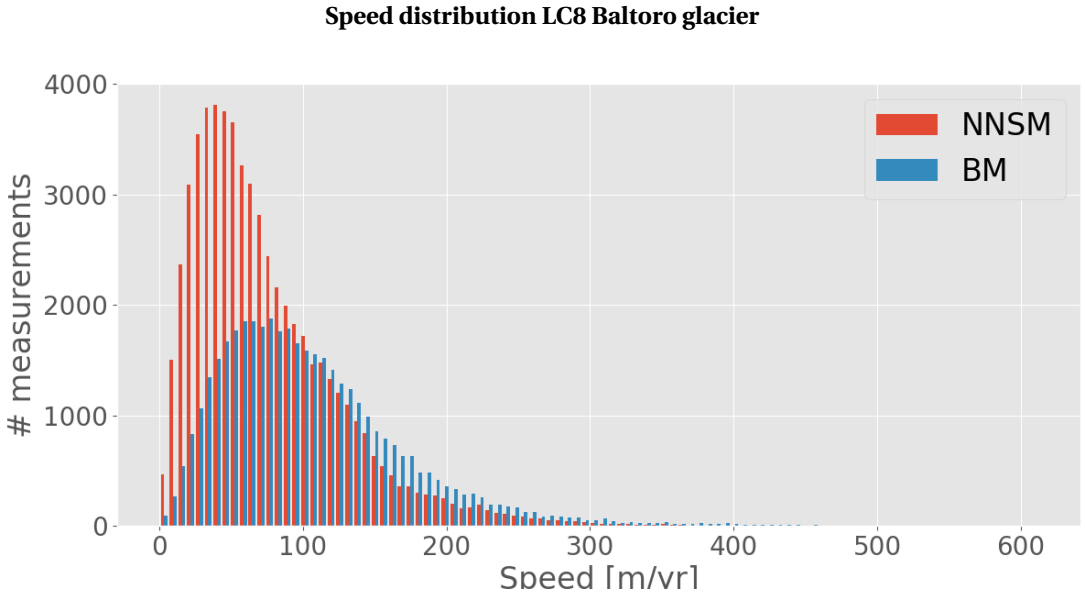


Figure 5.3: Speed distribution of Landsat 8 NN-stacking method and base method velocities of the Baltoro region.

5.1.2. External validation

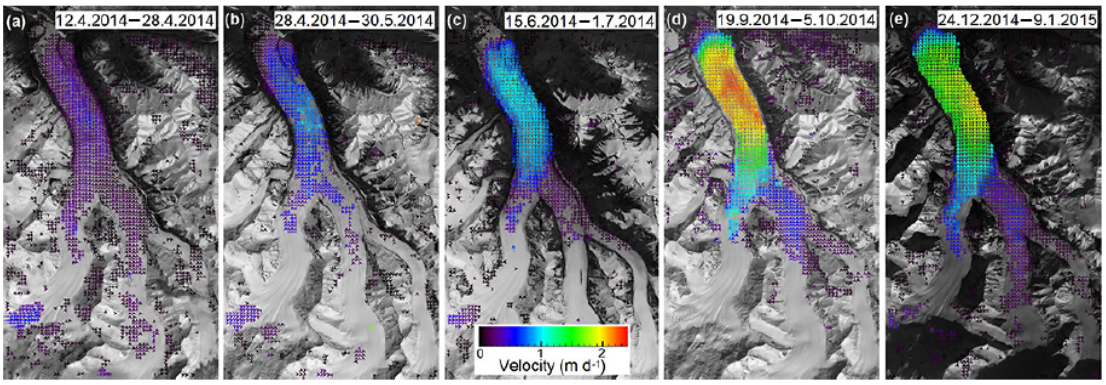
Results from [36] shows a surge of the Kyagar glacier using Landsat 8 data and feature tracking.

This is shown in figure 5.4a. The velocity of the glacier ranges from 200 m/yr to 800 m/yr at the peak of the surge and reaches peak surge velocity in 5 months. Figure 5.4b shows the velocity field for two equal periods. The velocity field from 2014-09-19 should show a velocity of around 800 m/yr, but the maximum velocity of the feature tracking was set at 500 m/yr. It is a good thing however that the resulting velocity field is filtered out and not replaced by a faulty velocity field. The velocity fields from 2014-04-12 are harder to compare, but both shows velocities of around 200 m/yr. This result provides external validation for the algorithm. Precise external validation can be provided by comparing velocity profiles from the Baltoro glacier.

The velocity profile is the velocity along the center line, see figure **??**, of a glacier. Higher on the glacier the velocity is larger and decreases towards the end of the glacier. This is depicted in figure 5.6a, which shows the velocity and altitude along the distance of the glacier. This velocity profile is created from feature tracking applied on ERS-1 and ERS-2 radar data and covers periods from 1996 to 2006. To validate results from both the base method and NN-stacking method the temporally closest good velocity profile, from 2001/02/27 - 2002/03/31, is selected from Landsat 7 and presented in figure 5.6b.

Comparing the Landsat 7 SLC-on velocity profile from figure 6.5b to 6.4a we see that the velocities from the NN-stacking method and base method are more variable than the results from ERS-1, ERS-2. This is both due to the co-registration and Imgraft errors amplified by the small temporal baseline as well as the large resolution (30 m) of the Landsat 7 images. The general trend of velocity profile is similar to the reference profiles from figure 5.6a, with a velocity somewhere in the middle of the different

a. Kyagar glacial surge from [36]



b. Kyagar glacial surge from NNSM

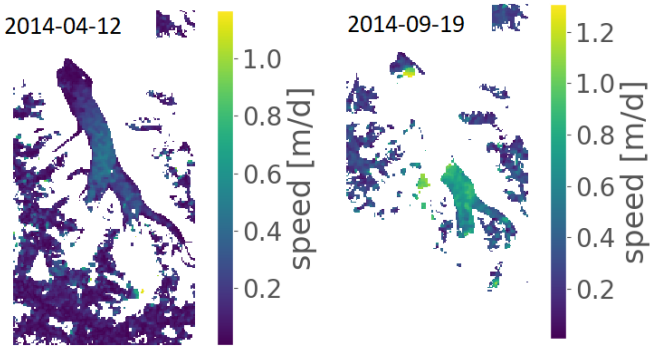


Figure 5.4: Velocity field of the Kyagar glacier from different time periods in 2013-2014, dates shown in the top of the images. In this period a glacial surge can be seen.

periods. In both profiles the peak velocity around 30 km from terminus is visible. The velocity profile from the NN-stacking method is a bit smoother compared to the profile from the base method. that is, large velocity variations are not altering from pixel to pixel as is the case for the base method velocity, but rather grouped along larger areas of pixels.

The NN-stacking method is thus both validated from comparison with the base method as from external validation. The NNSM shows a tendency to have lower velocities than the base method. This does not seem to be true from comparison with external validation, but as the time period from both results is different this is hard to verify. The next results section compares the performance between the NNSM and BM.

Center line of Baltoro glacier.

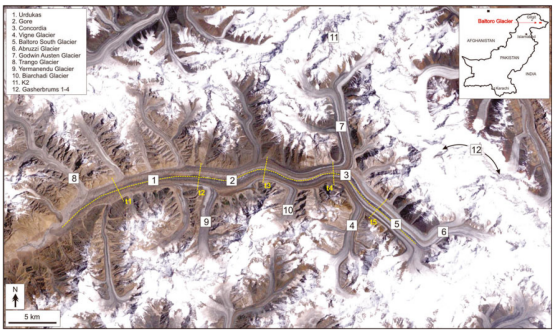
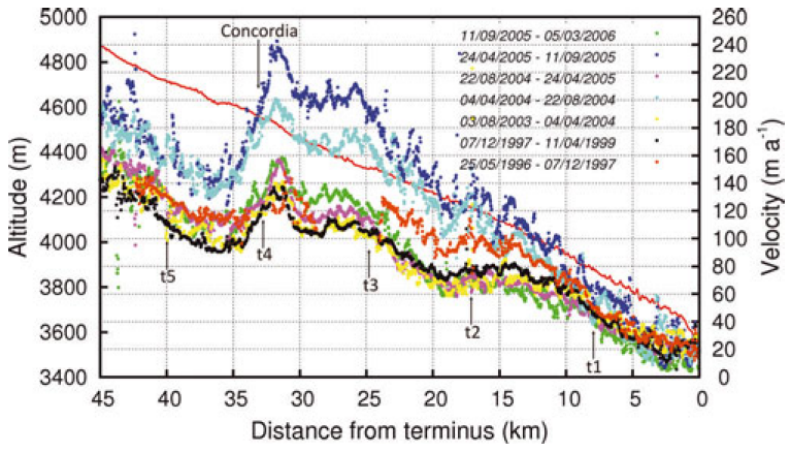


Figure 5.5: Center line of Baltoro glacier. This is the line running trough the middle of the glacier from the terminus (end) to the top of the glacier.

a. Velocity profile Baltoro glacier



b. Velocity profile Baltoro glacier Landsat 7

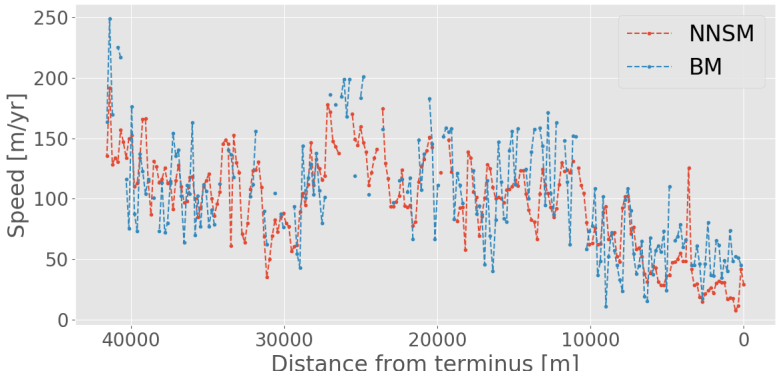


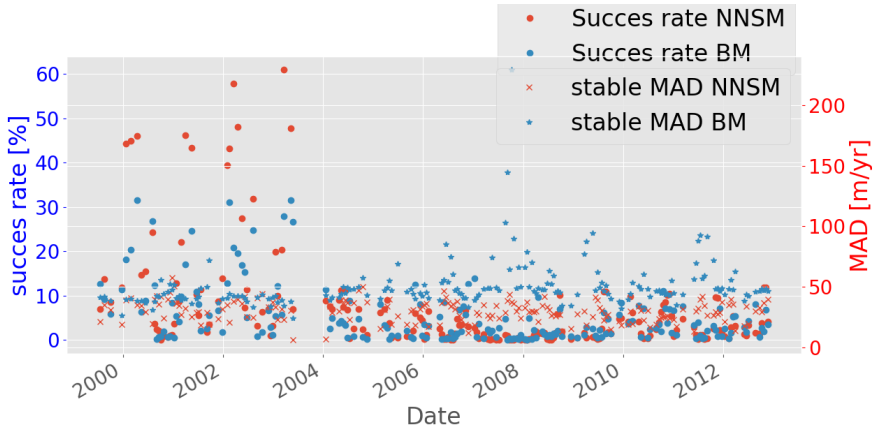
Figure 5.6: Velocity profiles of the Baltoro glacier. (a) Altitude and velocity profile from [31], velocities for different periods from 1996 - 2006. (b) Velocity profile from NNSM and BM Landsat 7 SLC-on data from 2002/02/27 - 2002/03/31.

5.2. Performance results

5.2.1. Success rate & stable MAD

The performance results are summarized by the statistics presented in table 5.2. Figure **??**a shows the success rate and stable MAD of both the NNSM and the BM over time for the Landsat 7 time series. Figure **??**b shows the difference of the success rate and stable MAD between the NNSM and BM over time. A positive success rate difference means that the success rate is higher for the NN-stacking method. A negative stable MAD means that the stable MAD is lower for the NN-stacking method, which is good.

a. Absolute performance Landsat 7



b. Performance difference of Landsat 8 BM and NNSM

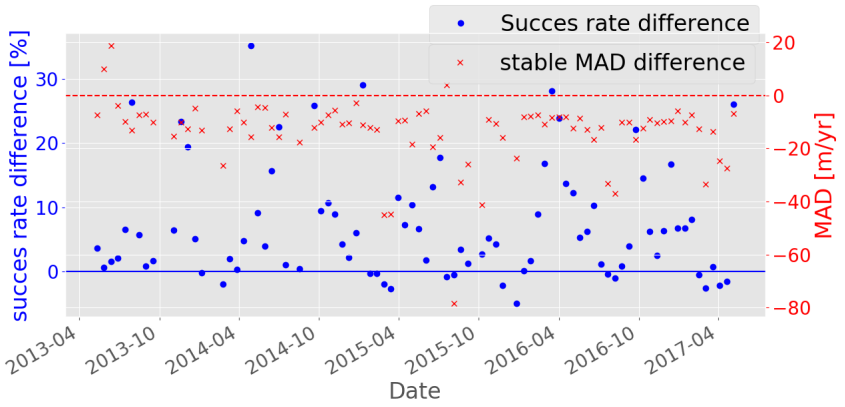


Figure 5.7: Performance plots of the Landsat 8 results for the Karakoram region. The left axis shows the success rate of the velocity field, the right axis the MAD on stable ground. The epoch of a velocity field is the epoch of the slave base pair image used for the velocity field. (a) shows the absolute performance per epoch for both the NN-stacking method and base method. (b) shows the difference between both methods.

Table 5.2 compares the average performance of all the consecutive velocity fields that make up the velocity times series per Landsat type between the NN-stacking method and the base method. This is expressed as the success rate, which is the percentage of all pixels on the glacier mask which are not masked. The stable MAD, e.g the error of the velocity on stable ground. The SR difference: the total percentage of pixels that have a lower or higher success rate compared to the other method. Thus, a 5% SR difference for the BM means that 5% of all glacier pixels have a temporal success rate lower than the base method. It is essentially a quantification of the distribution of the success rate difference around the blue line in figure 5.7b. Finally the average dispersion is the average error of the pixel velocity estimate using the NN-stacking method.

The results show an an increase of success rate of the NN-stacking method results over the base

Performance statistics

Landsat	Method	Average success rate	Average stable MAD [m/yr]	SR difference [%]	Average dispersion [m/yr]
5	base method	7	54.8	15.5	-
5	NNSM	12.1	42.7	30.7	133.4
7	base method	5.2	50.3	20.4	-
7	NNSM	7.1	29.8	24.8	100.4
8	base method	12.4	59.8	5.8	-
8	NNSM	19.6	48.2	38.6	85.7

Table 5.2: Averages of the success rate,stable MAD, SR difference and dispersion per sensor type for all time series velocity fields.

method for all Landsat types. The success rate is low for all Landsat types in this region, but especially low for Landsat 7 results because of the data gaps due to the failure of the scan line corrector. This can be seen in figure 5.7a, where a data gap can be seen in 2003-2004 after which the success rate decreases. The scenes after 2003-May-31 show poor performance in total and a robust time series cannot be created for Landsat 7 SLC-off scenes in this research. As the results now show that results from Landsat 7 ETM+ SLC-off and Landsat 5 TM scenes are not usable for time series research (in this research), the results shall focus on Landsat 8.

The stable MAD is smaller for all NN-stacking method velocity fields. The stable MAD of 48 m/yr of the Landsat 8 NN-stacking method compared to the 59 m/yr of the base method velocity fields indicates that the errors left in the velocity fields of the NN-stacking methods are smaller than those of the base method. The average dispersion, is the lowest for Landsat 8 and larger than the stable MAD. The dispersion also indicates the size of the velocity errors, but per pixel velocity estimate. As it is larger than the stable MAD, this indicates that the stacks of velocities from which the median velocity is estimated contains some faulty displacements increasing the magnitude of the dispersion and possibly offsetting the velocity estimate from the true velocity.

Temporal success rate distribution Landsat 8

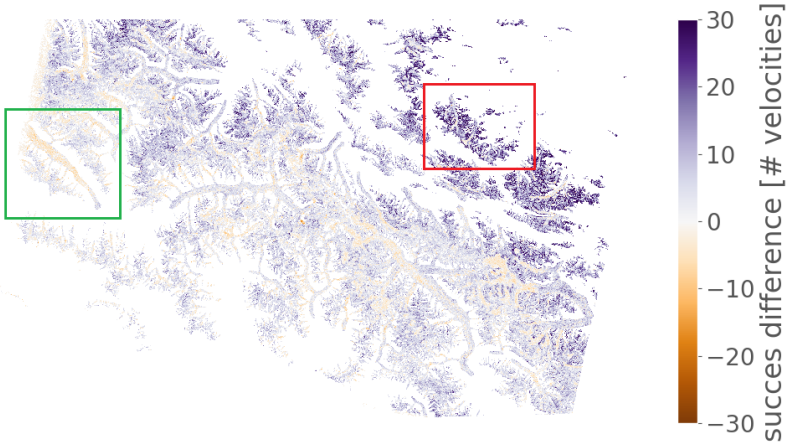


Figure 5.8: Spatio-temporal success rate difference between NN-stacking method and base method on glaciers in the Karakoram region.

As was noted in the previous section, the success rate is defined for all points within the glacier mask but not all points included in the glacier mask are flowing glaciers. Some also include nearby mountains which have snow cover all year. Furthermore is it assumed that the increase in success rate in due to correct velocities and not because of the addition of faulty velocities. The only large scale indication that the results present that the NN-stacking method does not see an increase in success rate due to faulty displacements is that the stable MAD also decreases. About the distribution of the success rate,

figure 5.8 shows the spatial distribution of the increase of success rate for Landsat 8.

The distribution shows that the NN-stacking method acts as a filter for some glaciers, as indicated by the green area in figure 5.8, where the success rate is lower. This glacier often showed faulty displacements using the base method. On the other hand, the success rate is increased due to glacier masked areas indicated by the red area in figure 5.8. This area contains little flowing glaciers, but contributes to the increase in success rate because it is marked as glacier area.

It is thus hard to prove the increased success rate (coverage) of the NN-stacking method over the base method on a large scale. To prove the increased performance of the NN-stacking method over the base method the next section also presents and compares the flow field stability of both methods.

5.2.2. Flow direction stability

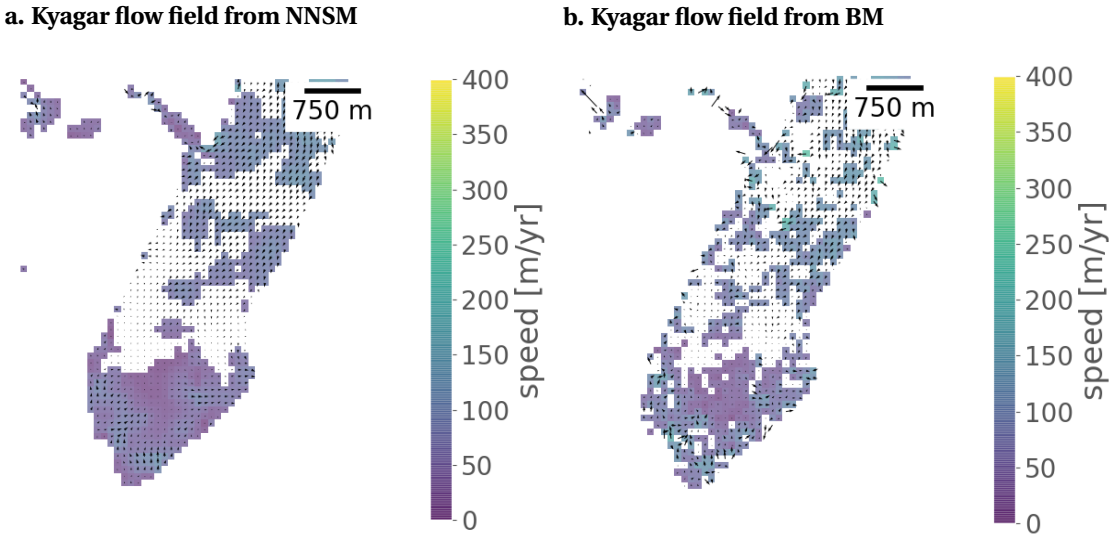


Figure 5.9: Flow field of the Kyagar glacier. The size of the direction arrows correspond with the flow speed. (a) shows the flow field of the NNSM. (b) shows the the flow field of the BM.

The stability of the flow direction is another performance indicator. Figure 5.9 shows the flow of the Kyagar glacier for the Landsat 8 velocity field from 2016/02/13-2016/02/29 from the base method and NN-stacking method. This figure is a zoomed in area of the Kyagar glacier and the vectors show the flow direction, while the size of the vectors correspond with the speed of the flow. The masked areas that show flow vectors, but not speed are areas filtered out by the temporal filter.

The RMSE between the NNSM and BM velocity is 33.6 m/yr for this zoomed in area of a flowing glacier, indicating that the RMSE from table 5.1 which includes the non-flowing glacier mask areas around the glaciers is a good indication of the RMSE on glaciers.

Two things can be noticed from the flow fields. First of all, the flow field from the NNSM is locally smoother than the base method flow field. The general smoothness of the flow can be expressed by the order parameter S, the isotropy of the flow along a director. The director of the flow is defined as a flow in the SSW (-110°) from the + X axis. The S parameter is 0.81 for the flow of the NN-stacking method and 0.74 for the base method flow indicating that the flow of the NNSM is a little bit more isotropic. Secondly at the end of the glacier, indicated by the red box in figure 5.9a, the speed shows a sudden increase and the flow is opposite to the rest of the glacier for both the BM and NNSM. As this is the glacier terminus, the speed is expected to be zero. This shows that errors dominate in this stagnant area and this corresponds with a different flow. On faster flowing glaciers, where the flow is well defined, the flow

could thus be used as an indication of performance.

The final performance indicator is the dispersion of the NN-stacking method velocity estimate.

5.2.3. Dispersion

The dispersion for the Landsat 8 and Landsat 7 SLC-on velocity field is presented in figure 5.10. The dispersion is the MAD of a velocity estimate of a pixel using the NN-stacking method. This also includes the pixels neighbors. Thus, when all faulty displacements are filtered out, the dispersion should indicate the errors due to the co-registration, orthorectification and precision of the Imgraft correlator and the velocity difference between the 9 neighboring pixels. For large velocity gradients the dispersion is thus expected to be larger. For low velocity gradients the magnitude of the dispersion is expected to be around the magnitude of the stable MAD.

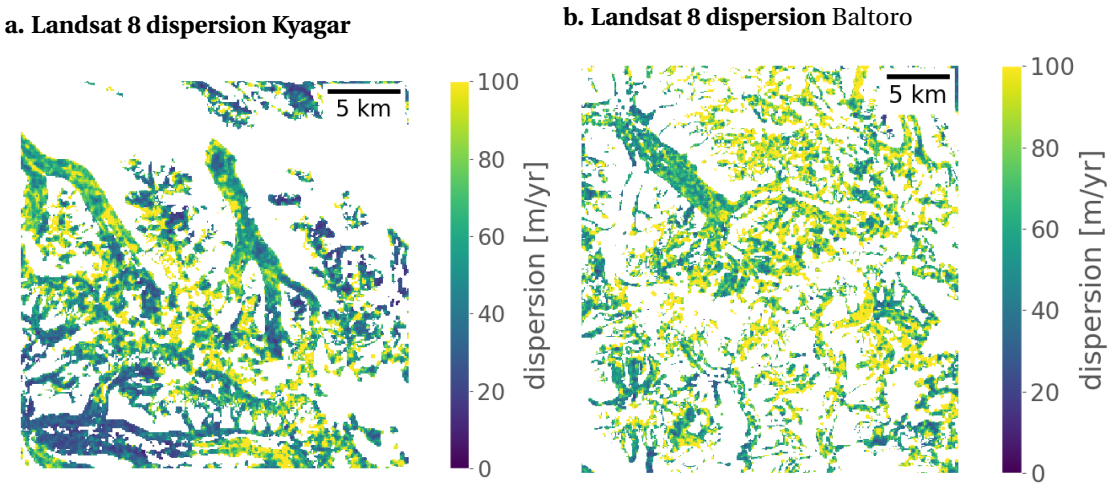


Figure 5.10: Velocity dispersion of NN-stacking method around the Baltoro glacier for (a) Landsat 8 (b) Landsat 7 SLC-on.

The dispersion is about ~50-70 m/yr on the Kyagar and Upper Baltoro glacier. This is larger than the stable MAD, but could be due to combined co-registration and Imgraft errors of only 0.75-1.75 meter over 16 days. This is much lower than expected from error propagation. Combined with the lower stable MAD from the NNSM compared to the BM, this indicates that the NN-stacking method decreases the error in the error propagation and the feature tracking method reduces co-registration errors.

At the areas around the glaciers, indicated in figure 5.2 by the red area, the dispersion is often larger than the dispersion on the glacier while the velocity is stable and low. This indicates that the velocity estimates still include faulty displacements. Thus the stack filtering before the velocity estimate in the NNSM could still be improved or the dispersion could be used as a filter afterwards.

As the results validate the NNSM and show an increased performance from the NNSM compared to the base method we can check if a robust dense time series is now improved by the NN-stacking method compared to the base method.

5.3. Time series

Figure 6.9 shows the velocity time series of a point on the center line of the Upper Baltoro glacier, around 40 km from the terminus. A velocity of about 140 m/yr with variations of 25 m/yr can be expected for this point. The x-axis shows the date and the date of the velocity is expressed as the epoch of the slave image of the velocity base pair. Plotted are both the NN-stacking method and base method. Velocities linked by a line are consecutive velocities.

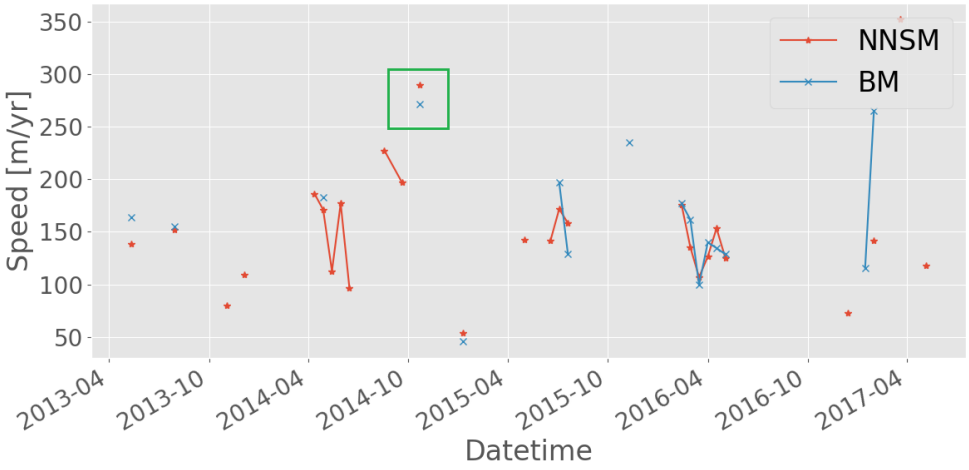


Figure 5.11: Velocity time series of point on the center line of the Baltoro glacier, around 40 km from the terminus.

The time series show more velocities for the NN-stacking method and fewer outliers. That is, fewer velocities outisde of the expected velocity range. However, it also shows large jumps between consecutive velocities as can be seen around 2014-05. In total the time series is clearly not dense enough to estimate seasonal variations and it is not sure if the results are accurate enough to measure the expected seasonal variations of about 50 m/yr. The outlier indicated by the green area in figure 5.11 is studied by selecting the two base Landsat 8 scenes from which one is presented in figure 5.12. We see that the outlier is created by thin clouds in one of the base pair scenes. While the results of the NN-stacking method thus show in improvement in filtering faulty velocities not all the results are yet correct.

This result is more visible when plotting a time series from the Baltoro velocity profile, as shown in figure 5.13. This shows the time series of the Baltoro velocity profile from the section "External Validation" for both the NN-stacking method and the base method. The expected velocity profile is similar to the results from figure 5.6 assuming the profile did not change completely in the period 2006-2014. That means a velocity of about 150 m/yr at the beginning of the glacier, around 45 km from the terminus and 0 m/yr at the terminus. Results from both methods still show outliers. The expected velocity profile is much more visible from the NNSM result, as much more outliers have been filtered out by this method. It is also clear that the current tuning and setup is not yet perfect as an automated approach should filter out almost all outliers, but many outliers still remain visible, even in the results from the NNSM method.

5.4. Summary

The results for the Karakorum region show most importantly that Landsat 5 and Landsat 7 SLC-off velocities are unsuitable for consecutive time series. This is due to the scan line artifacts from the Landsat 5 TM scanner and the low success rate and errors induced by the scan line failure from the Landsat ETM+ scanner.

The validation of the NN-stacking method comes from two sources. The first is comparison with the base method (single velocity field), where the NNSM shows to have a bias towards lower velocities, but the velocity fields on the glaciers show a similar distribution. The second validation is the comparison with results from other research using both Landsat and radar data. It shows that the algorithm is able to create correct velocities, but local variability is induced due to errors that are enhanced by the short temporal baseline and the rather large resolution of the selected Landsat 7 images.

Landsat 8 base scene 14/10/21

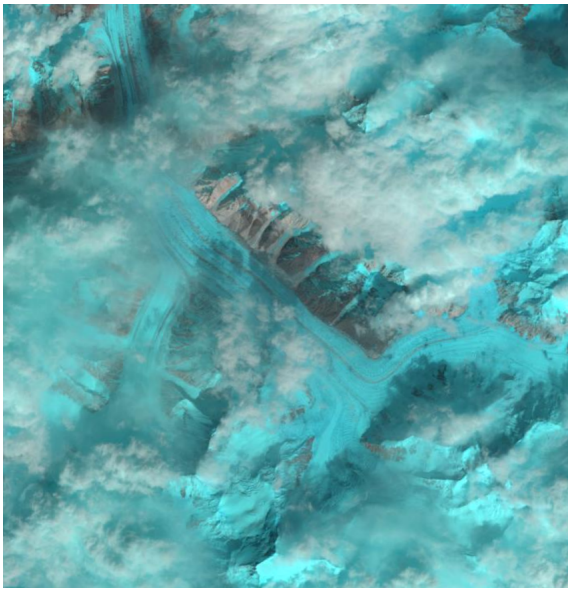
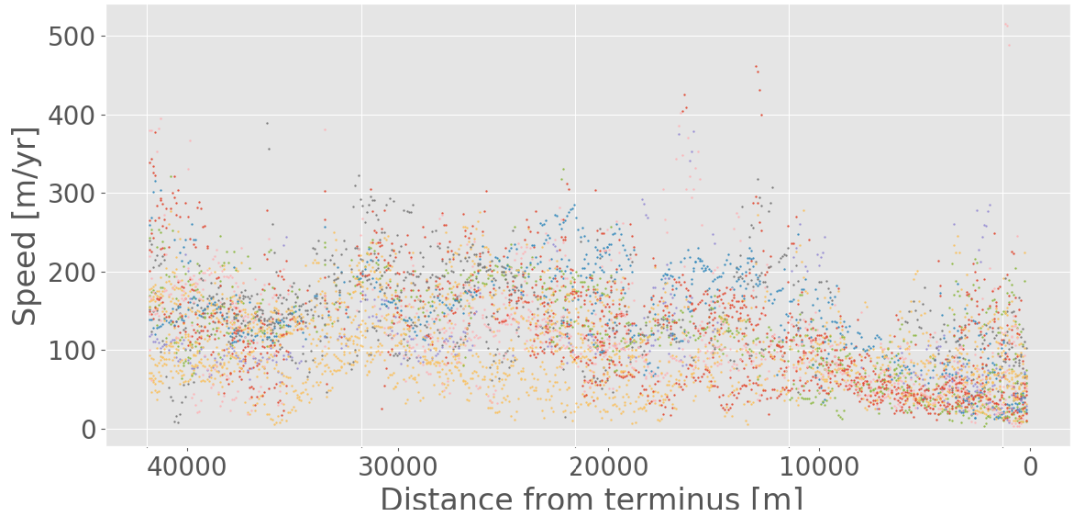


Figure 5.12: Landsat 8 image from 2014/04/21 of domain around the Upper Baltoro glacier. Notice the thin clouds and cloud shadows in the region.

The performance of the NNSM is better than the base method for all sensor types. The results from the NN-stacking method have a smoother flow, reduced errors and a higher success rate. Furthermore it is accompanied by an error indication, the dispersion. The NNSM was shown to act as a filter for erroneous velocities.

Time series from both methods still have outliers however, although the NNSM does reduce outliers. The times series are too sparse and not robust enough to show short term seasonal velocity variations in the Karakoram. The result is good enough though to show larger velocity variations like the glacial surge of the Kyagar glacier.

a. NNSM Baltoro velocity profiles Landsat 8



b. BM Baltoro velocity profiles Landsat 8

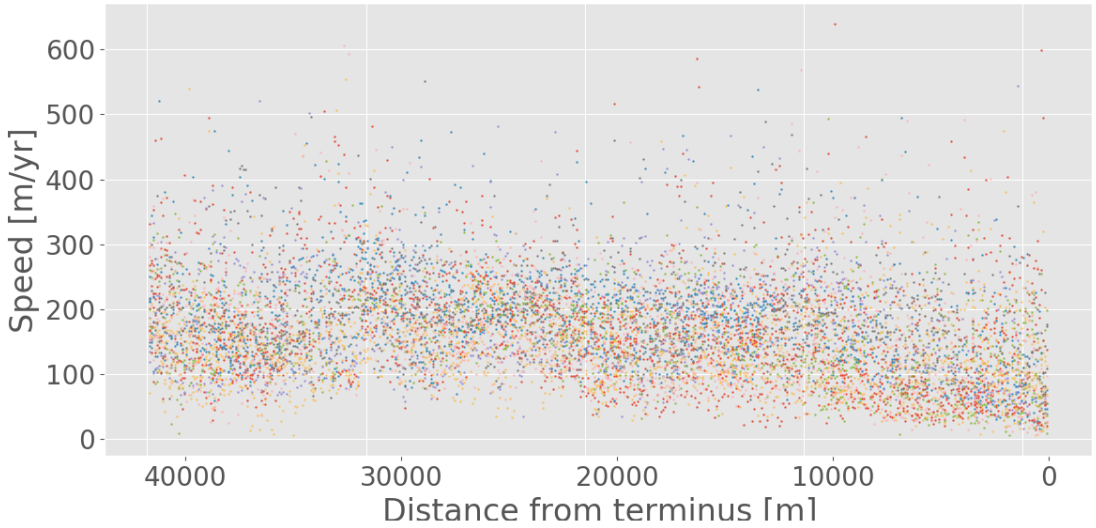


Figure 5.13: Velocity profile of Baltoro glacier for all velocity fields in the Landsat 8 time series from (a) NN-stacking method (b) base method.

6

Results: Everest

In total 6753 velocity fields have been produced for this region, 2167 Landsat 5 velocity fields, 3056 Landsat 7 velocity fields and 1530 Landsat 8 velocity fields. The total processing time was about 10 days, with 7 days for only the Landsat 8 results. These velocity fields are combined using the proposed method to create a consecutive dense time series of 187 velocity fields for Landsat 5, 200 velocity fields for Landsat 7 and 84 velocity fields for Landsat 8.The Everest region has stable and slow glaciers: the speed often does not exceed 70 m/yr. The two studied glaciers, Ngozumpa and Kangshung have velocities ranging from 0-50 m/yr.
This chapter validates results from the NN-stacking method using the base method and external validation, compares the performance of the base method and NN-stacking method and presents time series results.

6.1. Validation of NN-stacking method

6.1.1. Base method validation

The results for Landsat 5 and Landsat 7 SLC-off velocity fields are dominated by scan line artifacts and the velocity results show artificial North-South velocity variations due to these artifacts. The velocity fields for the Ngozumpa and Kanshung glacier are presented for Landsat 8 (2016/10/18-2016/11/03) and Landsat 7 SLC-on (2000/10/14-2000/11/13) in figure 6.1 and 6.2.

These velocity fields show that both glaciers are slow moving and the velocity on the glacier is in the expected range of < 70 m/yr. However, the area in the red box north of the Ngozumpa glacier in figure 6.1b shows velocities larger than 70 m/yr. This area is the start of the glacier and a icy steep slope with highly variable velocities which are not really part of the glacier and has no external validation. The red box around the glacier terminus in figure 6.1 is expected to have almost no velocity (see the next section, figure 6.4. However, these results show a velocity of ~20-50 m/yr.

The Landsat 7 SLC-on results do not show the scan line artifacts as noticed in Landsat 5 TM velocities. Both NN-stacking method results are smoothed compared to the base method results. This is especially true for the Landsat 8 base method results, where the increased resolution and relatively small reference window size gives rise to local variability of the speed on the glaciers. Both the base method and NN-stacking method results seem to be unable to obtain zero velocity on stable ground or at the glacier terminus. This is confirmed by the average stable velocity which is 14.3 m/yr for the Landsat 8 velocity field of figure 6.1. Notice that this means that the co-registration bias (offset) of the two scenes to create the velocity field is in the order of 0.05-0.1 meter in the x and y direction.

The accuracy of the NN-stacking method is expressed as the pixel-wise RMSE between the NN-stacking method velocity and base method velocity in table 6.1 and the speed distribution is depicted in figure 6.3 for the Landsat 8 Ngozumpa domain. The RMSE is expressed in m/yr and includes all the

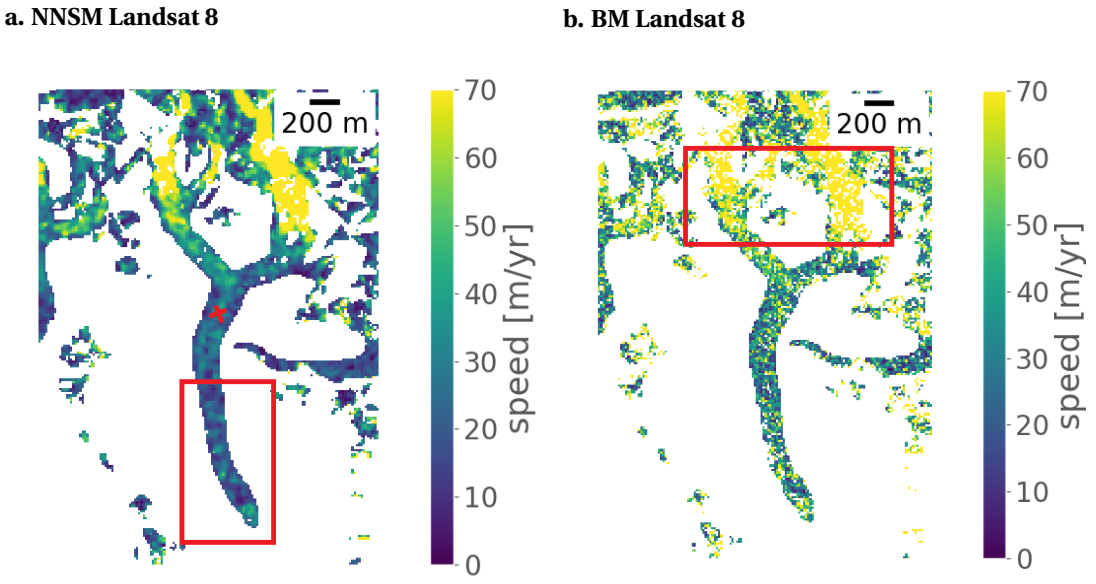


Figure 6.1: Landsat 8 velocity field of domain around the Ngozumpa glacier for both the NN-stacking method and base method . (a) shows the NN-stacking method results, (b) shows the base method results.

glacier masked areas, thus also the icy slopes around the glacier which might offset the RMSE results.

Landsat	RMSE Ngozumpa [m/yr]	RMSE Kanshung [m/yr]
LT5	34.1	27.3
LE7	31.2	29.6
LC8	32.7	28.8

Table 6.1: Root mean square error between the NN-stacking method velocities and the base method velocities for all Landsat types and the Ngozumpa and Kangshung glacier.

Analysis of the RMSE shows that the error has a magnitude equal to the expected velocity on glaciers in the Everest region [32]. The speed distribution shows that is not only due to differences in the variance of both methods, but that the velocity of the NN-stacking method is biased towards a lower velocity. Furthermore, the speed distribution shows that most or all of the velocities larger than 100 m/yr are filtered out using the NN-stacking method. From visual inspection of figures 6.1 and 6.2 we can see that this filtering and bias towards lower velocity is due to the filtering and smoothing effect of the NN-stacking method for stable areas around the glacier as well as on the glacier. Also, because both velocity fields have a short temporal baseline of 16 and 32 days respectively and the glaciers are slow moving, the effect of co-registration and Imgraft errors are visible in both the base method results as local variability of the velocity.

It was assumed that the results of the base method are good and the results of the NN-stacking method can be validated from the base method. The large local variability of the base method however makes it hard to validate the NN-stacking method results. External validation provides better validation options.

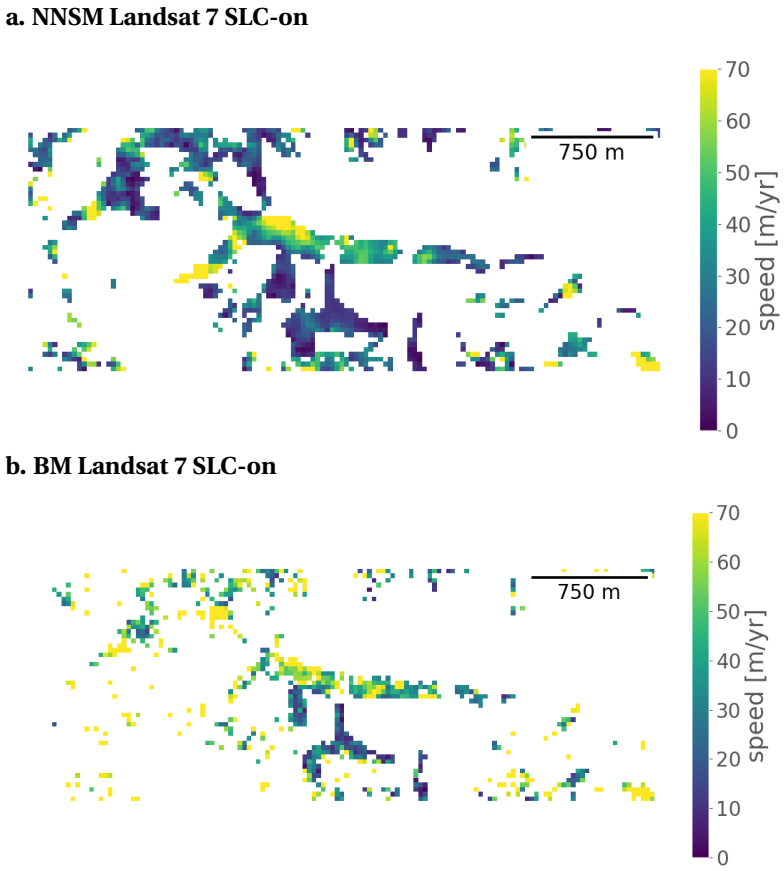


Figure 6.2: Landsat 7 SLC-on velocity field of domain around the Kanshung glacier for both the NN-stacking method and base method . (a) shows the NN-stacking method, (b) shows the base method.

6.1.2. External validation

Results from [32] show a maximum velocity of 45-50 m/yr for the Ngozumpa and Kangshung glacier, with a large stagnant area for the Ngozumpa glacier 6 km from the terminus, as shown in figure 6.4b and indicated by the red box in figure 6.1a. The data used for these results is from ERS SAR data, and the shown results use a combination of interferometry and feature tracking to create velocity fields. The feature tracking results use a 1 year temporal baseline and the dates are presented in the bottom right of the velocity fields.

To validate the results from the base method and NN-stacking method, the closest LE7 SLC-on velocity field with enough velocities on the Kangshung glacier is selected, from 2000/12/21-2001/01/20. The velocity profile for the base method and NN-stacking method is shown in figure 6.5a. A velocity profile obtained from Landsat 8 results is also shown in 6.5b, from 2016/12/01-2016/12/17. The velocity field used for comparison with figure 6.4a is the same as the the velocity fields presented in figure 6.2.

Comparing the Landsat 7 SLC-on velocity profile from figure 6.5a to 6.4b we see that while the NN-stacking method velocity profile is already more stable than the base method profile, the velocity is still too high compared to the external results. A small downward slope is visible towards the terminus in the NN-stacking method result as expected, but the inconsistent results from 1500 m to 500 from terminus make it difficult to see. It should be noted that this might be because the time period is different for the velocity profiles and is actually a correct result. It also shows that the temporal filter with a threshold of 40 m/yr is too large for regions with slow moving glaciers. The velocity fields from figure 6.2 also show that the velocity is too high for the base method velocity and better but still too high for the NN-stacking method velocity. It is also important to see that local variations of velocity are prone to be larger because

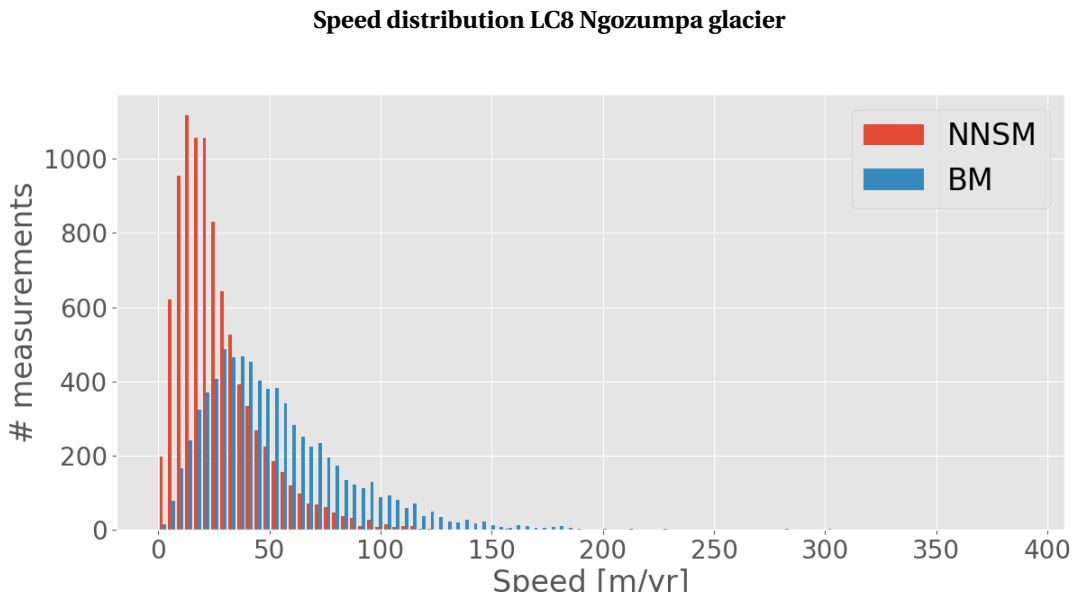


Figure 6.3: Speed distribution of Landsat 8 NN-stacking method and base method velocities of the Ngozumpa region.

of the low resolution on these small glaciers. The results only consists of a couple of pixels or even a single pixel towards the terminus of the glacier.

The velocity profile from Landsat 8 shows the same increase in stability for the NN-stacking method compared to the base method velocity profile. Also, assuming the velocity profile did not change much in the period from 2001 to 2017 the Landsat 8 NN-stacking method result shows good similarity to the result in 6.4b. The velocity high on the glacier is around 35 m/yr and drops towards the terminus of the glaciers to around 15 m/yr. However, the NN-stacking method result still shows local variations that are large compared to the total change of velocity over the glacier.

The external validation shows that the bias towards lower velocities of the NN-stacking method is good, because it is closer to the true velocity of the external results and the spatial stability of the velocity is better for the NN-stacking method results. The performance results of the NN-stacking method are presented in the next section.

6.2. Performance results

6.2.1. Success rate & stable MAD

The performance results are summarized by the statistics presented in table 6.2. Results from the Karakoram region showed that these key statistics and especially the success rate can be deceiving. The performance of the Landsat 8 time series is shown in figure 6.6

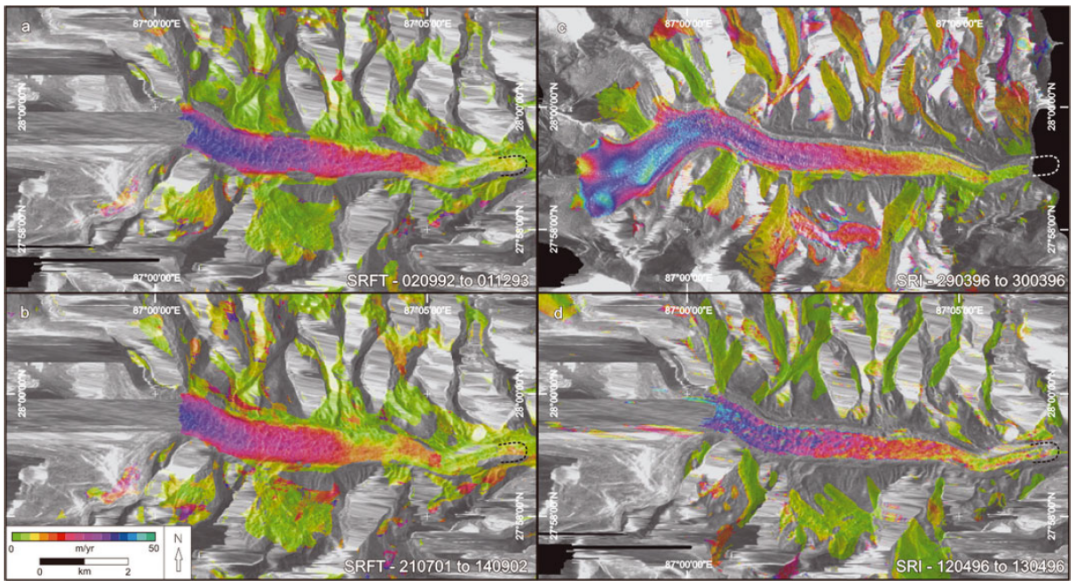
Performance statistics

Landsat	Method	Average success rate	Average MAD [m/yr]	SR difference [%]	Average dispersion [m/yr]
5	base method	10	48.6	3.5	-
5	NN-stacking method	22.1	39.2	45.3	121
7	base method	8.6	48.2	6.5	-
7	NN-stacking method	21.7	25.9	41.6	113.1
8	base method	25.5	48.7	6.9	-
8	NN-stacking method	35.2	27.5	41.5	75.3

Table 6.2: Averages of the success rate,stable MAD, SR difference and dispersion per sensor type for all time series velocity fields.

Table 6.2 shows an increase of success rate of the NN-stacking method results over the base method

a. Velocity fields



b. Velocity profiles

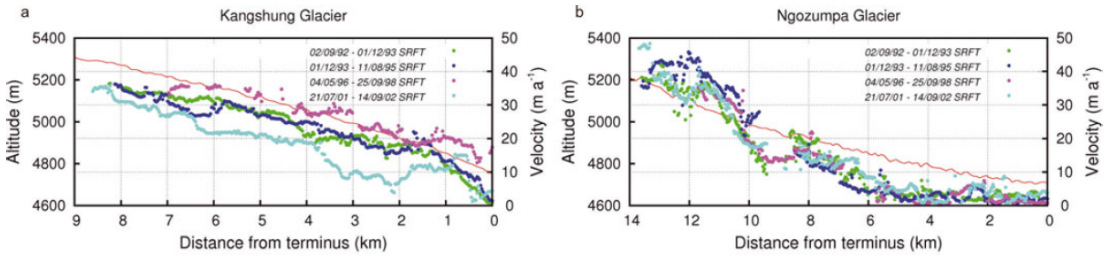


Figure 6.4: External velocity fields and profiles for the Kangshung and Nogzumpa glacier from [32]. The velocity fields show results from feature tracking (a,c) and interferometry (b,d) for the Kangshung glacier. The velocity profiles show feature tracking results for (a) the Kangshung glacier and (b) the Ngozumpa glacier.

for all Landsat types, but the increase is the smallest (~10%) for Landsat 8 results. The stable MAD on the other hand is smaller for all NN-stacking method velocity fields. The stable MAD of 27.5 m/yr of the Landsat 8 NN-stacking method compared to the 48.7 m/yr of the base method velocity field indicates that the errors left in the velocity fields of the NN-stacking methods are smaller than those of the base method. The average dispersion, variance of the velocity estimate of a pixel, is the lowest for Landsat 8 and about three times as large as the stable MAD. The dispersion also indicates the size of the velocity errors. As it is larger than the stable MAD, this indicates that the stacks of velocities from which the median velocity is estimated contains some faulty displacements increasing the magnitude of the dispersion and possibly offsetting the velocity estimate from the true velocity. Finally the SR difference shows the distribution of the success rate difference, that is the distribution of the success rate difference around the blue zero line in figure 6.6b. The 6.9 % of the Landsat 8 BM result thus says that of all possible glacier pixels that could have a velocity, 6.9% of these pixels has fewer velocity results than the NN-stacking method.

The success rate is defined for all points within the glacier mask but not all points included in the glacier mask are flowing glaciers. Some also include nearby mountains which have snow cover all year. Thus, the increase of the success rate could be due to an increase of results from the areas around the glaciers. From figure 6.10 we can see that the increase of success rate is evenly spread over the glacier mask and thus also true for flowing glaciers. The stable MAD of the Landsat 5 velocity fields shows that

a. Velocity profile Kangshung glacier Landsat 7 **b. Velocity profile Kangshung glacier Landsat 8**

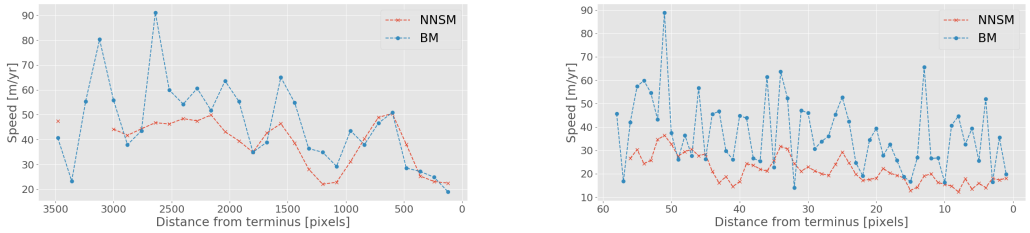


Figure 6.5: Velocity profile of the Kangshung glacier from (a) Landsat 7 SLC-on and (b) Landsat 8 results. Results compare the profile from the base method and NN-stacking method

the scan line artifacts increase the error in the results. It is also noted that the larger dispersion of Landsat 5-7 results compared to Landsat 8 has to be because of the difference in selected spectral bands and possibly radiometric performance, indicating that the use of spectral band 8 with 15 meter resolution for Landsat 7 might be a better choice.

A very interesting result is seen from the changes of the success rate and stable MAD over time. Figure 6.6 shows that the success rate is highest in January-March and lowest in July-September. The MAD on the other hand increases in the period July-September. The summer period of the Everest region is monsoon period, cloud cover is more frequent thus lowering the success rate and increasing the stable MAD, while surface changes due to precipitation also influence results. The resolution of the results and strict filtering prevents the algorithm from finding results between the clouds and thus the NN-stacking method does not increase performance much in this period.

6.2.2. Flow direction stability

The stability of the flow direction is another performance indicator. Figure 6.7 shows the flow of the Kangshung glacier for the Landsat 8 results from the base method and NN-stacking method. This figure is a zoomed in area of the Kangshung glacier. The flow of the Kangshung glacier is from West to East (left to right in the figure), and larger in the west at the beginning of the glacier.

Visual inspection of figure 6.7 shows that the the flow of both methods is generally in the correct direction. This can be checked by the order parameter S. The assigned director (reference flow) is + 45 from the +x axis (NW). The order parameter S is 0.62 for the NN-stacking method and 0.66 for the base method. This indicates that both flows follow this reference direction reasonably well, but are also not complete isotropic. The local smoothness of the flow is however not expressed by this order parameter and is better for the NN-stacking method.

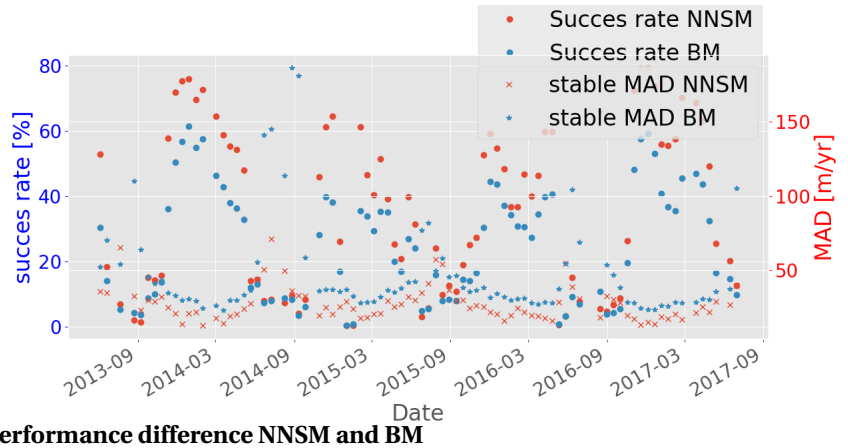
The RMSE between the NN-stacking method and base method is 23.5 m/yr for this zoomed in glacier region. Thus, the RMSE is lower when only using flowing glacier data, indicating that the RMSE is influenced by ice covered non-flowing glacier regions. The S parameter shows that both methods have a similar performance on general flow of the Kangshung glacier, but tells us nothing about the local flow which is better for the NN-stacking method as we can see from visual inspection of figure 6.7.

6.2.3. Dispersion

The dispersion for the Landsat 8 and Landsat 7 SLC-on velocity field is presented in figure

The dispersion shows, similar to the stable MAD, the size of the errors of a velocity field. That is, when all faulty displacements are filtered out it should show the errors due to the co-registration, orthorectification and precision of the Imgraft correlator and thus should show errors similar of size to the stable MAD. The dispersion is about 30 m/yr on the Ngozumpa glacier for both Landsat results corresponding with the magnitude of the stable MAD.

a. Absolute performance of NNSM and BM



b. Performance difference NNSM and BM

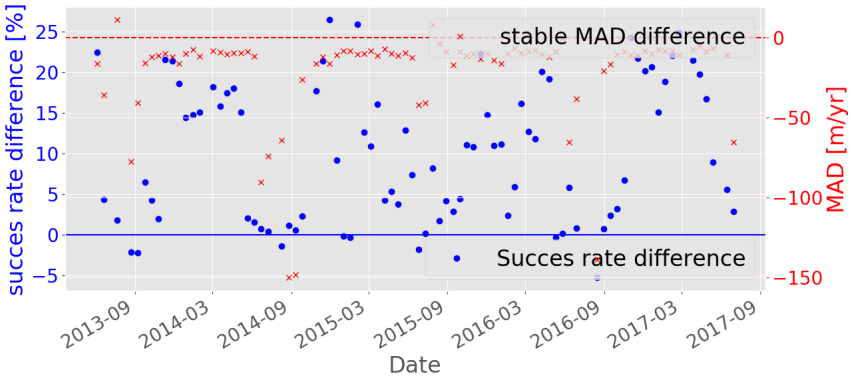


Figure 6.6: Performance plots of the Landsat 8 results for the Everest region. The left axis shows the success rate of the velocity field, the right axis the MAD on stable ground. The epoch of a velocity field is the epoch of the second base pair image used for the velocity field. (a) shows the absolute performance per epoch for both the NN-stacking method and base method. (b) shows the difference between both methods.

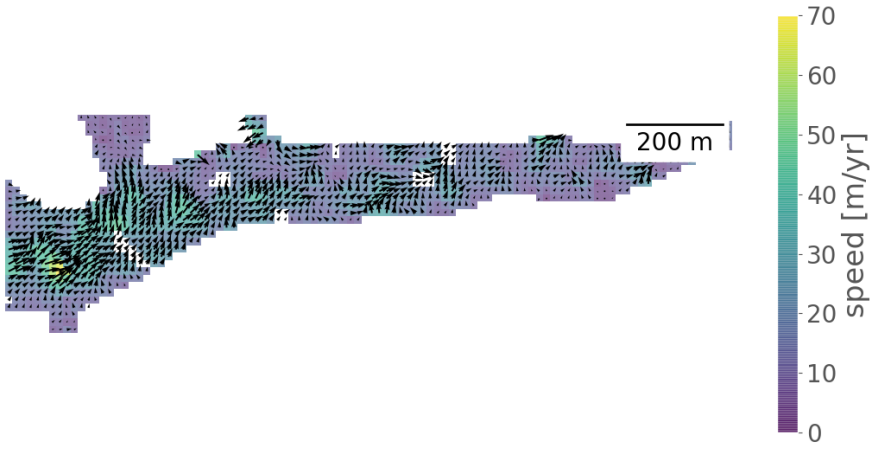
The red box north of the Ngozumpa glacier in figure 6.8a. shows a large dispersion, and this area coincides with the unexpected large velocity in the indicated area of figure 6.1b. The dispersion of the indicated area near the terminus (south) of the Ngozumpa glacier also corresponds with a large unexpected velocity. This thus indicates that these velocities could be wrong and should be filtered out or improved.

6.3. Time series

Figure 6.9 shows the velocity time series of a point on the center line of the Ngozumpa glacier, around 12 km from the terminus indicated by the red cross in figure 6.1a. The x-axis shows the date and the date of the velocity is expressed as the epoch of the slave image of the velocity base pair. Plotted are both the NN-stacking method and base method. Velocities linked by a line are consecutive velocities.

From 6.4 a velocity of about 30-40 m/yr is expected not varying more than 10 m/yr between seasons. The results show relatively large jumps in velocity between epochs for both the NN-stacking method and base method. The average velocity of the NN-stacking method is closer to the actual expected velocity and has fewer large variations between epochs, while the velocities from the base method are consistently too high. It is again clear that the temporal filter threshold, which should filter out large jumps in the time series is too large for this slow moving glacier or the filter does not use enough tem-

a. Kangshung flow field from NNSM



b. Kangshung flow field from BM

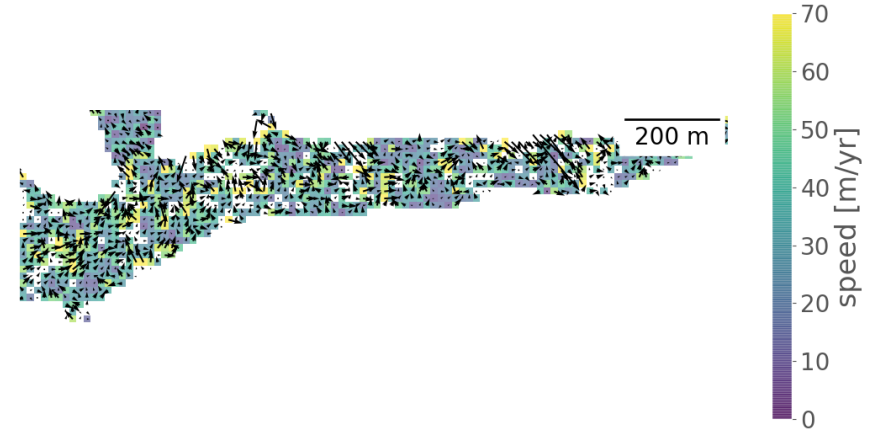


Figure 6.7: Flow field of the Kangshung glacier. The size of the arrows correspond with the flow speed. (a) shows the flow field from the NN-stacking method. (b) shows the flow field from the base method.

poral information. Also, even though the NN-stacking method time series is more stable than the base method time series, the result is not accurate enough to measure seasonal changes for this slow moving glacier.

Figure 6.10 shows the temporal success rate difference for the glacier masked area in the Everest region. This glacier mask area consists of actual flowing glaciers, as for example the Ngzumpa and Kangshung glaciers indicated by the red areas, but also of snowy/icy areas around the glaciers (indicated by the blue area). The figure shows that the increase in success rate of the NN-stacking method is evenly distributed over the glacier mask and thus is also due to an increase on flowing glaciers like the Ngozumpa and Kangshung glaciers.

6.4. Summary

The results from the Everest region again show the failure to obtain good velocity fields for Landsat 4-5 TM scenes and Landsat 7 ETM+ SLC-off scenes. The RMSE between the NN-stacking method and base method velocity is in the order of the expected velocity (30 - 40 m/yr) of this region and biased towards lower velocities for the NN-stacking method. There is a constant velocity offset of 10-20 m/yr, thus zero velocity on stable ground is hard to achieve. Comparing the results to external validation however the

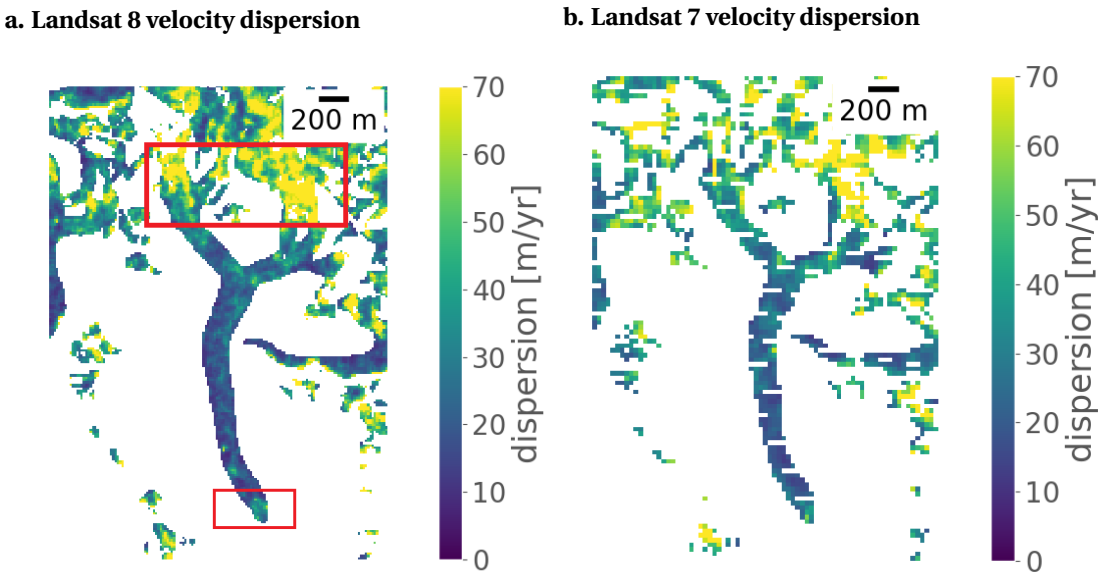


Figure 6.8: Velocity dispersion of NN-stacking method around the Ngozumpa glacier for (a) Landsat 8 (b) Landsat 7 SLC-on.

NN-stacking method results are closer to the expected velocities in the region and produce a better velocity profile than the base method, especially for Landsat 8. It was shown that the lower resolution of Landsat 7 bands 4-5 shows disadvantages over the higher resolution of Landsat 8 band 8.

The performance results show an increased performance of the NN-stacking method over the base method for all Landsat missions and especially for Landsat 8 results. It also shows a low success rate and higher MAD in the summer when snow and cloud cover affect the results from the feature tracking and almost no results are available. The local smoothness of the flow is better for the NN-stacking method results, but the overall isotropy of the flow is similar for both results from the NN-stacking method and the base method. It is also noted that the dispersion of the results is often as large as the velocity on the glaciers and a high dispersion is visible on regions where poor results are expected (icy slopes of nearby hills). However, the dispersion on glaciers is often of the same magnitude as the stable MAD showing the there are little or no outliers left on glaciers from the velocity estimates using the NN-stacking method.

The time series show a large variability between epochs compared to the expected seasonal variations. Measuring seasonal changes for slow moving glaciers is therefor impossible using the current algorithm setup.

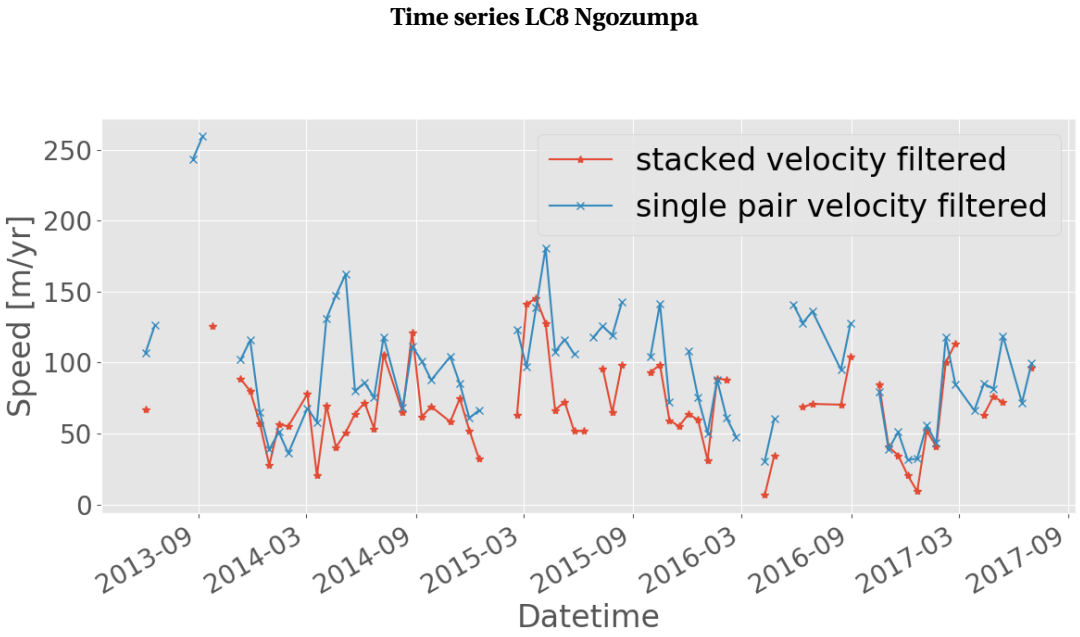


Figure 6.9: Velocity time series of point on the center line of Ngozumpa glacier, around 12 km from the terminus.

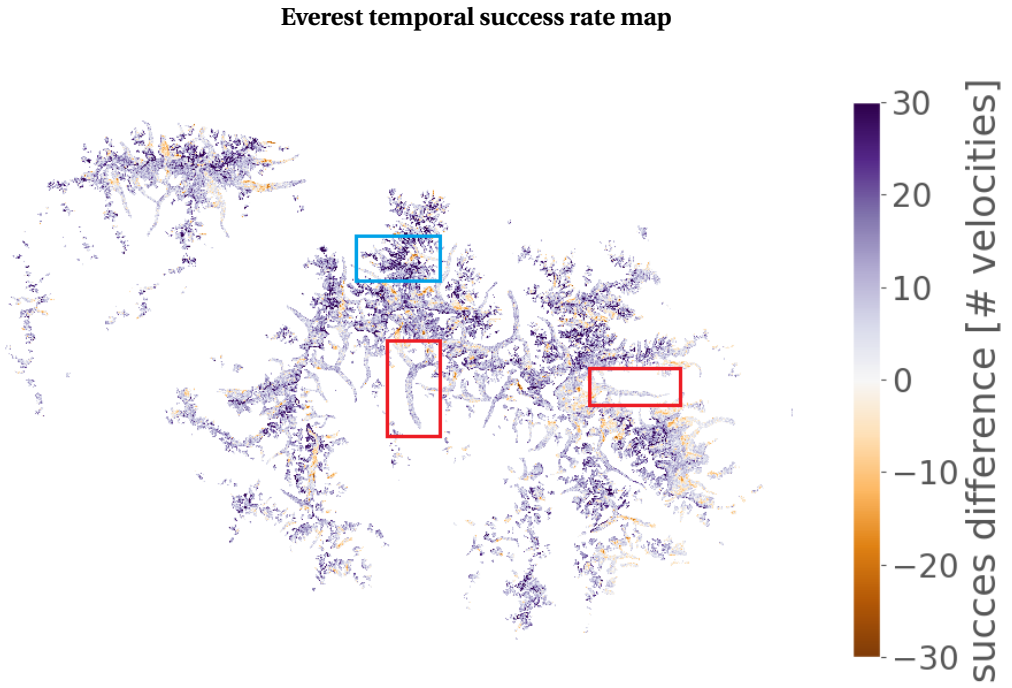


Figure 6.10: Spatio-temporal success rate difference between NN-stacking method and base method of glaciers in the Everest region.

7

Conclusions & Recommendations

In this research a new method for creating complete (consecutive) time series of glacier velocities created from Landsat images using the RIFT algorithm is proposed and tested in the Himalayas. This chapter answers the research questions linked to this method and briefly discusses these conclusions. The second part of this chapter provides recommendations on future use of this method and the analysis of glacier velocity time series.

7.1. Summary, Conclusions & Discussion

To answer the main research question the literature and method are summarized, related sub questions are answered and conclusions are discussed. To provide good conclusions it should be kept in mind that the goal of this research is to improve methods for automatically creating dense glacial velocity time series for large areas. This requires internal validation as well as good temporal and spatial coverage.

The questions: **Using the current (BM) algorithm, could a precise time series be made spanning all Landsat missions?** and **What are the errors in the velocity fields, can they be quantified and how do they appear in the velocity fields?** are answered from a summary of the literature research on Landsat and feature tracking.

7.1.1. Landsat

This research uses Landsat optical images (scenes) as input for glacier velocity calculations. This is because the Landsat database is large and is thus suitable for automated large area research. The creation of time series of glacier velocities requires temporal accuracy of the optical Landsat measurements.That is, pixels from Landsat scenes from different epochs should be represent the same location on earth. A study of the Landsat missions and sensors shows that the performance of Landsat missions before Landsat 4 is not suitable for time series. This is because of the poor radiometric performance and coregistration of scenes between epochs as well as a different scene reference system (WRS-1). Missions from Landsat 4 and on using the TM, ETM+ and OLI-TIRS sensors have suitable spectral bands and radiometric performance and all use the WRS-2 (path/row) reference system. The Landsat 8 OLI-TIRS scenes are especially suitable, as the radiometric performance is much better. This improved radiometric performance and good resolution improves results from feature tracking on snow and ice. The increased managing and pre-processing of the USGS of Landsat scenes from Landsat 4 and later also makes them suitable. The quality of these Landsat scenes is divided in tiers. Scenes from the best tiers and same path and row are temporally coregistered in this pre-processing within 12 meter RRMSE, have information on errors and quality indicators for snow, shadows and clouds and are not fully clouded. Errors from Landsat scenes are thus only coregistration errors, which are provided with the Landsat scenes. Orthorectification errors (errors due to faulty projections) are small when using the same path and row and are not easy to model. One problem with the Landsat database however is the failure of the scan

line correction from Landsat 7 ETM+ scenes at 31-May-2003. The results are horizontal lines of data gaps in scenes after this epoch with up to 25% of the data missing.

7.1.2. Feature Tracking

Velocity of glacial flow is calculated from two Landsat scenes (image pair) using repeat image feature tracking (RIFT) algorithms. These RIFT algorithms calculate the displacement of features in the x and y-direction using cross-correlation functions.These cross-correlations can be calculated in different ways. Based on from research from [17] orientation correlation (OC) is selected as the most suitable for the Himalayas. As this calculation takes place in the Fourier domain, it is also faster than other methods which is useful for large area research. Feature tracking algorithms also otput a singal-to-noise ratio (SNR). This provides an indication of how valid the displacement is and can be used to filter out faulty results due to for example clouds, snow cover or shadows. A good threshold for SNR filtering is a known problem using these algorithms as a threshold that is too low leaves faulty displacements and too high filters out valid displacement. Success of feature tracking displacements depends on snow cover, shadows, clouds and the time between the images used as input of the feature tracking (temporal baseline).

We can conclude from this that, while the Landsat database is large, much cannot be used for time series analysis. The time spans are not consistent for older Landsat missions, and due to performance Landsat 1-3 images are not useful for this research and time series and Landsat 7 images only partly. Errors are due to the precision of feature tracking and coregistration and are known. Finally it is noted that from literature the selection of a good SNR threshold for filtering is a known problem.

7.1.3. Methods & performance testing

Answers to the sub questions **What is the performance of a method and how can it be tested?** and **How can combinations of image pairs be used to create time series and how should it improve performance?** are provided here.

The research compares two methods of creating time series. One method, the base method (BM), uses just one filtered velocity field between two consecutive Landsat scenes and creates a time series from this. These Landsat scenes are the master and slave image and is called the base pair. The proposed method uses the same base pair of Landsat images, but also combinations of filtered velocity fields that add up to the same velocity field as calculated from the base pair images. It thus creates a stack of velocity fields that should be the same same except for errors. The median of these velocities then estimates a single velocity field that is equal to the base pair velocity field. Three types of methods for acquiring velocity fields from this stack of values were proposed: the single pixel stacking method (SPSM), spatial filtering stacking method (SFSM) and nearest neighbor stacking method (NNSM) .The NNSM was selected as the best performing method of these three, based on visual inspection of the results and comparing performances of the methods. This result however was only from one path/row and only for Landsat 5 TM scenes. The NNSM performed better mostly because it filtered out speckle and the stack size a velocity was estimated from is much larger. The minimal stack size of the SPSM and SFSM method was only 5 velocities. Estimating the velocity from such a small set is prone to be affected by outliers even though the median estimator is robust. Also, as the period between consecutive scenes is sometimes large for Landsat 5 and 7 scenes, a maximum temporal baseline of 2 years was set for combinations. It was concluded that the selected NNSM method does have a smoothing effect and is in some way self validating due to the use of neighbors in the estimate.

To assess the quality of a velocity field several performance indicators have been used. These are the spatial success rate, MAD on stable ground (variability of results on stable ground, should be close to zero), the temporal success rate (how many velocities are there in an time series, a directional order parameter that indicates how well the velocity results are oriented along the flow on glaciers and the smoothness of the time series. The MAD and spatial success rate are familiar performance indicators in the field of glacier velocities, however they are only focused on spatial performance. The temporal success rate gives an extra indication of where a time series shows good performance or what method shows better performance as a function of space and time. The smoothness of time series was introduced to serve as a temporal alternative of the stable MAD, but the proposed 1-lag autocorrelation that

outputs values between 0 and 1 as an indicator of performance does not work for sparse time series. Finally to study not only the speed of the flow but also the flow direction, the orientation order parameter S was introduced as another spatial performance indicator. Although this does output good information on the isotropy of the flow, the flow direction is only useful on glaciers that have a flow speed much larger than the expected errors and cannot output a global indication of performance as it depends on deviations from local directors of flow. It is thus hard gain much information from this indicator that is of use for large scale performance. It might be useful however for local filtering on faster moving glaciers.

7.1.4. Single velocity fields and Imgraft

This research used the Imgraft toolbox to calculate the displacements and thus velocity fields from preprocessed Landsat image pairs. These image pairs can be a single spectral band or a combination of spectral bands. These different spectral bands can be combined in a pre-processing step by using principal component analysis to create a single image. After this a Gaussian high-pass filter is applied to both images. The feature tracking parameters are the size of the search window, which is calculated from the expected maximum velocity in the region, the size of the reference window and spacing between correlation windows. Results of a velocity field in this research thus depend on six parameters: spectral bands, size of the high pass filter, the three feature tracking parameters and the temporal baseline between the two images. The other values selected for the parameters are summarized in table 7.1.

Region	Sensor	band(s)	high pass filter (pixels)	reference window (pixels)	window spacing (pixels)	maximum velocity (m\yr)
Karakorum	LT5	4,5	1000 meter	12	6	500
Karakorum	LE7	4,5	1000 meter	12	6	500
Karakorum	LC8	8	1000 meter	12	6	500
Everest	LT5	4,5	1000 meter	12	6	500
Everest	LE7	4,5	1000 meter	12	6	500
Everest	LC8	8	1000 meter	12	6	500

Table 7.1: Final parameter settings per test area and sensor type.

The selected values of the feature tracking parameters are based on a combined testing, where both the maximum velocity and reference window were changed. However, one of both parameters might have had more weight changing the results. It is suspected that the selection of the maximum velocity (and thus a large search window) has a large effect on the variability as larger outliers are allowed and the correlation is more likely to find a faulty displacement. Also research from [9] shows that for reference window sizes above 12 pixels the difference in variability of the results is small. The size of the high pass filter kernel was chosen based on personal and visual testing and is thus not very reliable. An important mistake made in this research was the use of a 12 pixel reference window size for Landsat 8 scenes. As the resolution used for these scenes is only 15 meters, the size of the reference window should have been 24 pixels (360 meter).

An important conclusion is the failure of all methods to create good velocity fields for Landsat 7 SLC-off scenes and Landsat 5 TM scenes. For Landsat 7 SLC-on results, the method of filling the data gaps with white noise as to improve correlation results was concluded to give worse results than leaving them masked. As the orientation correlation of Imgraft does return displacement results even with the data gaps and the limited 8 bit white noise range cannot guarantee good randomness this backs this conclusion. The velocity fields of Landsat SLC-off scenes however, can not not be trusted using either method without manual inspection. The Landsat 5 TM velocity results have large offsets due to scan line artifacts from the TM scanner. These are magnified by the short temporal baselines.

7.1.5. Comparison NNSM and base method and validation of NNSM

The remaining conclusions answer the sub questions **Which combination of feature tracking techniques, parameters and algorithm parameters gives the best result in the areas of research?** and **What are the results of validation of the new method and application on the test areas?**.

The NNSM is validated from two sources. The first is comparison with the base method. This concludes that the NN-stacking method results are biased towards lower velocities. External validation shows that this bias actually gives better results for slow moving glaciers. The external validation also validates results from both the NN-stacking method and the base method and shows that the NNSM is spatially more stable.

The performance of the NNSM is better than the performance of the base method. The success rate is often similar or higher. In some regions however the success rate is lower for the NNSM. This is mostly because of the filtering effect of the NN-stacking method and in these areas faulty velocities are filtered. The stable MAD is lower for velocity fields from the NNSM indicating that large errors due to co-registration, Imgraft precision and orthorectification and faulty velocities are more often filtered out by the NNSM. The flow field from the NN-stacking method is locally smoother and more isotropic. Also, as the velocities from the NNSM come with a local error indication, the dispersion. The dispersion shows that the NNSM parameter tuning is not yet perfect, but does provide a good indication of robustness of results.

In conclusion, results from the NN-stacking method are validated from two sources and have a better performance than the base method.

7.1.6. Time series

The final parameters for the time series algorithm using the NNSM are presented in table 7.2. The number of combinations and maximum temporal baseline make sure the stack size is large enough for improving estimates of the true velocity, while not using temporally spaced scenes where the correlation is mostly lost. The minimal stack size is selected as 20 values, but this parameter setting is not tested or tuned and the effect of this parameter on the performance is unknown. The temporal filter requires smoothness from the time series, but the filter threshold is also not tuned. It has been seen that a static temporal filter threshold does not function for different glacial areas.

The presented time series show less erratic behaviour for velocities calculated using the NNSM, but this is not quantified as the proposed performance indicator to express this can not deal with the sparseness of the time series. The variability in the time series is too high to accurately measure seasonal velocity fluctuations using the current settings of the algorithm. This is both due to leftover outliers and errors from co-registration and Imgraft.

Region	Sub methods	Nr combinations	Max temporal baseline [#]	Minimal stack size	SNR threshold [-]	Temporal filter threshold $[\Delta m \backslash yr]$
All	NNSM	30	730 days	20	4	70

Table 7.2: Final parameter settings for the NN-stacking method.

7.1.7. Final conclusion

The result of this research is an implemented method for automated glacial velocity time series creation using the NN-stacking method. It also tests the complete algorithm and provides parameter settings for the Himalayas region. The research tried to answer the question: *Is it possible to make a velocity time series from the complete optical Landsat mission for the Asian Highlands where the performance, e.g the number of results and validity of results, is increased by using combinations of image pairs opposed to using single image pairs?*

The main conclusion from this research is that the proposed method is not yet suitable for creating dense time series automatically on a large scale for the Himalayas. Using the current parameter and algorithm setup, the co-registration and Imgraft errors are too large to provide a precision that is better than the seasonal variability in the Himalayas. Furthermore the current setup does not automatically filter out all outliers due to clouds, shadows and snow. It was however already able to measure a glacial surge and does show similar results to velocity profiles from other researches.

Using the NN-stacking method however, the performance is better than a time series of single velocity fields. The coverage is similar to the base method, but it filters more outliers and provides an indication of error from the velocity dispersion. Furthermore the method is not limited to optical velocity fields, but can be applied to any series of velocity fields. The drawback of the NNSM is the large computational effort compared to the base method.

7.2. Recommendations

The method for creating time series and the implemented algorithm show good performance for coherence, validation and stability of time series results compared to the base method. Future automated research by use of this algorithm could thus be useful. To be useful however some problems in the existing algorithm need to be solved and some fundamental problems should be taken into consideration.

7.2.1. Landsat

The two case studies show that the Landsat 4-5 TM is unreliable for large scale time series results due to the scanning artifacts and this research was not able to obtain good results for Landsat 7 ETM+ SLC-off results. Results from [35] show that time series results for Landsat 7 are possible however. Thus it would be best to only use Landsat 7 and Landsat 8.

The pre-processing from USGS for Landsat scenes has also improved the last years. Next to the co-registration improvements, Landsat scenes now include quality indicators. This includes cloud masks, indicators for intensity saturated pixels and snow/ice indicators. This information could be used as an extra condition for filtering of the results.

7.2.2. Co-registration

Dense time series means that velocity fields often have short temporal baseline and only a small displacement. Thus the errors (co-registration, feature tracking precision and orthorectification) will be large compared to the displacement. The proposed method solves this by using redundancy, but no extra co-registration was used in this research. The use of improved co-registration and a more precise feature tracking method is thus recommended for implementation into the algorithm.

7.2.3. Repeat image feature tracking

The repeat image feature tracking algorithm focuses on scenes from only one path/row. However, Landsat scenes do overlap, i.e measurements from different orbits do overlap. It was shown from multiple sources [3], [38] [11] that using accurate co-registration and orthorectification can be used to map these neighboring scenes to one scene. The result is a higher frequency of optical images from which a denser time series can be created.

Another drawback of the RIFT algorithm and the proposed method in particular is the large computational cost of the correlation algorithms. The calculation of the results in this research took about 9-10 days. This is while the analysis was run in parallel on a supercomputer. A novel method for calculating glacial velocities was already studied in [4]. This research uses an optical flow algorithm used normally in computer vision, turbulence studies and self driving cars. The computational performance of the optical flow algorithm is much better, thus very suitable for automated large scale analysis.

7.2.4. Filtering

Filtering of the results is an important step. The two filters used for the base method are filtering on the SNR and the rolling mean filter. It is however recommended to use a rolling median filter instead, but implementation of a 3D rolling median filter is difficult. This is because a rolling median filter is more useful for regions with dynamic glaciers. Along these two filters the use of the NNSM and the minimal amount of points from which the median is calculated, functions as an extra filter.

Literature on filtering commonly suggests very simple filters like a Hampel spatial filter, but with the main focus on the SNR filter. These filters rely on spatial validation and spatial variability to filter faulty

velocities, but the parameters for these kind of filters are hard to select such that they are suitable for large area research. Smarter algorithms like seed growing [11] are more suitable for large area research and it is recommended to investigate similar filtering methods. For example, noise is often grouped in areas of large variability. The use of for example k-median clustering to cluster areas on variance might be a filtering technique that does not require a difficult selection of filtering parameters for large areas.

7.2.5. Parameters

The sensitivity analysis in this research only consisted of investigating the SNR and the window size parameters. Previous research and testing methods on the window size parameters provide a good indication of the effects on changes to the window sizes and proposed settings for many areas around the world. The effects of the high pass filter however are less studied and while there is a proposed lower limit the effect of changes to the size of the high pass filter could be of interest. Furthermore, the NNSM provides two more parameters: the number of velocity field combinations and the minimal number of points (minimal stack size) from which to calculate the median. A sensitivity study is especially of interest for the minimal stack size in combination with the SNR threshold.

7.2.6. NNSM improvements

An easy improvement for the NNSM is the addition of another boundary condition before median estimation from the velocity stack. It has been noted many times that the SNR threshold is a poor filter as it is not always correct, especially for clouds. The NNSM could be improved by lowering the SNR threshold, but adding another boundary condition for valid velocities. This condition would be that the sum of the displacements should be zero. That is, for example, $d_{1-2} - d_{1-3} + d_{2-3} = 0$. If the residuals from this sum would be larger than a certain threshold, the combination of displacements could be removed from the stack. This is essentially checking the velocities in a stack for outliers to a reference velocity and requires that either d_{1-2} is correct or that a reference velocity field is available. If a reference velocity field is not available, an iterating procedure could be used to verify that d_{1-2} is correct. To do this, a low SNR threshold is used in the first iteration. This creates a reference velocity with a dispersion. Any values in the velocity stack that are outside the bounds dictated by this reference velocity and the dispersion are then filtered out and another velocity estimate provides a new reference velocity and dispersion. This process could be optimized to filter out all or most faulty velocities, thus improving the velocity estimate and reducing the error (dispersion) of an estimate.

7.2.7. Final recommendation

The complete task of creating large scale dense time series automatically is not easy. This is mainly due to the different characteristics of glacial areas and local climates. It essentially comes down to a complex optimization of success rate, stable MAD, dispersion and computational efficiency and all of this without validation. To solve this, this glacial research domain should become more closely related to computer vision. The NN-stacking method as presented in this research is a good first improvement. A major benefit of this method is that it is not limited to velocity fields from Landsat images, but can be used from any data source.

Bibliography

[1]

[2] Yushin Ahn and Ian M Howat. Efficient automated glacier surface velocity measurement from repeat images using multi-image/multichip and null exclusion feature tracking. *IEEE Transactions on Geoscience and Remote Sensing*, 49(8):2838–2846, 2011.

[3] Bas Altena and Andreas Kääb. Elevation change and improved velocity retrieval using orthorectified optical satellite data from different orbits. *Remote Sensing*, 9(3):300, 2017.

[4] Bas Altena and Andreas Kääb. Weekly glacier flow estimation from dense satellite time series using adapted optical flow technology. *Frontiers in Earth Science*, 5:53, 2017.

[5] Tim P Barnett, Jennifer C Adam, and Dennis P Lettenmaier. Potential impacts of a warming climate on water availability in snow-dominated regions. *Nature*, 438(7066):303–309, 2005.

[6] Paolo Berardino, Gianfranco Fornaro, Riccardo Lanari, and Eugenio Sansosti. A new algorithm for surface deformation monitoring based on small baseline differential sar interferograms. *IEEE Transactions on Geoscience and Remote Sensing*, 40(11):2375–2383, 2002.

[7] Gyanesh Chander, Brian L Markham, and Julia A Barsi. Revised landsat-5 thematic mapper radiometric calibration. *IEEE Geoscience and remote sensing letters*, 4(3):490–494, 2007.

[8] Misganu Debella-Gilo and Andreas Kääb. Measurement of surface displacement and deformation of mass movements using least squares matching of repeat high resolution satellite and aerial images. *Remote Sensing*, 4(1):43–67, 2012.

[9] Amaury Dehecq, Noel Gourmelen, and Emmanuel Trouvé. Deriving large-scale glacier velocities from a complete satellite archive: Application to the pamir–karakoram–himalaya. *Remote Sensing of Environment*, 162:55–66, 2015.

[10] ON Dhar and Shobha Nandargi. Hydrometeorological aspects of floods in india. In *Flood Problem and Management in South Asia*, pages 1–33. Springer, 2003.

[11] Mark Fahnestock, Ted Scambos, Twila Moon, Alex Gardner, Terry Haran, and Marin Klinger. Rapid large-area mapping of ice flow using landsat 8. *Remote Sensing of Environment*, 2015.

[12] Daniel Farinotti, Matthias Huss, Andreas Bauder, Martin Funk, and Martin Truffer. A method to estimate the ice volume and ice-thickness distribution of alpine glaciers. *Journal of Glaciology*, 55 (191):422–430, 2009.

[13] CB Field, VR Barros, DJ Dokken, KJ Mach, MD Mastrandrea, TE Bilir, M Chatterjee, KL Ebi, YO Estrada, RC Genova, et al. Ipcc, 2014: Climate change 2014: Impacts, adaptation, and vulnerability. part a: Global and sectoral aspects. contribution of working group ii to the fifth assessment report of the intergovernmental panel on climate change, 2014.

[14] AJ Fitch, Alexander Kadyrov, William J Christmas, and Josef Kittler. Orientation correlation. In *BMVC*, pages 1–10, 2002.

[15] Alex S Gardner, Geir Moholdt, J Graham Cogley, Bert Wouters, Anthony A Arendt, John Wahr, Etienne Berthier, Regine Hock, W Tad Pfeffer, Georg Kaser, et al. A reconciled estimate of glacier contributions to sea level rise: 2003 to 2009. *science*, 340(6134):852–857, 2013.

[16] T Haug, A Kääb, and P Skvarca. Monitoring ice shelf velocities from repeat modis and landsat data–a method study on the larsen c ice shelf, antarctic peninsula, and 10 other ice shelves around antarctica. *The Cryosphere*, 4(2):161–178, 2010.

[17] T Heid and A Kääb. Evaluation of existing image matching methods for deriving glacier surface displacements globally from optical satellite imagery. *Remote Sensing of Environment*, 118:339–355, 2012.

[18] T Heid and A Kääb. Repeat optical satellite images reveal widespread and long term decrease in land-terminating glacier speeds. *The Cryosphere*, 6(2):467–478, 2012.

[19] Dennis L Helder, Timothy A Ruggles, James D Dewald, and Sriharsha Madhavan. Landsat-5 thematic mapper reflective-band radiometric stability. *IEEE Transactions on Geoscience and Remote Sensing*, 42(12):2730–2746, 2004.

[20] Chengquan Huang, John RG Townshend, Shunlin Liang, Satya NV Kalluri, and Ruth S DeFries. Impact of sensor's point spread function on land cover characterization: assessment and deconvolution. *Remote Sensing of Environment*, 80(2):203–212, 2002.

[21] Seongsu Jeong and Ian M Howat. Performance of landsat 8 operational land imager for mapping ice sheet velocity. *Remote Sensing of Environment*, 170:90–101, 2015.

[22] G Kaser, JG Cogley, MB Dyurgerov, MF Meier, and A Ohmura. Mass balance of glaciers and ice caps: consensus estimates for 1961–2004. *Geophysical Research Letters*, 33(19), 2006.

[23] D Scott Lee, James C Storey, Michael J Choate, and Ronald W Hayes. Four years of landsat-7 on-orbit geometric calibration and performance. *IEEE Transactions on Geoscience and Remote Sensing*, 42(12):2786–2795, 2004.

[24] G. J.-M. C. Leysinger Vieli and G. H. Gudmundsson. On estimating length fluctuations of glaciers caused by changes in climatic forcing. *Journal of Geophysical Research: Earth Surface*, 109(F1): n/a–n/a, 2004. ISSN 2156-2202. doi: 10.1029/2003JF000027. URL `http://dx.doi.org/10.1029/2003JF000027`. F01007.

[25] Brian L Markham and John L Barker. Radiometric properties of us processed landsat mss data. *Remote Sensing of Environment*, 22(1):39–71, 1987.

[26] Brian L Markham, Md Obaidul Haque, Julia A Barsi, Esad Micijevic, Dennis L Helder, Kurtis J Thome, David Aaron, and Jeffrey S Czapla-Myers. Landsat-7 etm+: 12 years on-orbit reflective-band radiometric performance. *IEEE transactions on geoscience and remote sensing*, 50(5):2056–2062, 2012.

[27] RW McNabb, Regine Hock, Shad O'Neel, Lowell A Rasmussen, Y Ahn, M Braun, H Conway, S Herreid, I Joughin, WT Pfeffer, et al. Using surface velocities to calculate ice thickness and bed topography: a case study at columbia glacier, alaska, usa. *Journal of Glaciology*, 58(212):1151–1164, 2012.

[28] A Messerli and A Grinsted. Image georectification and feature tracking toolbox: Imgraft. *Geoscientific Instrumentation, Methods and Data Systems*, 4(1):23, 2015.

[29] Ron Morfitt, Julia Barsi, Raviv Levy, Brian Markham, Esad Micijevic, Lawrence Ong, Pat Scaramuzza, and Kelly Vanderwerff. Landsat-8 operational land imager (oli) radiometric performance on-orbit. *Remote Sensing*, 7(2):2208–2237, 2015.

[30] Remote Sensing Notes. Japan association of remote sensing/national space development agency of japan (nasda)/remote sensing technology center of japan (restec). *Asian Center for Research on Remote Sensing (ACRoRS) in Asian Institute of Technology (AIT)., JARS*, 1999.

[31] DJ Quincey, L Copland, C Mayer, M Bishop, A Luckman, and M Belo. Ice velocity and climate variations for baltoro glacier, pakistan. *Journal of Glaciology*, 55(194):1061–1071, 2009.

[32] DJ Quincey, A Luckman, and D Benn. Quantification of everest region glacier velocities between 1992 and 2002, using satellite radar interferometry and feature tracking. *Journal of Glaciology*, 55 (192):596–606, 2009.

[33] Eric Rignot and Pannir Kanagaratnam. Changes in the velocity structure of the greenland ice sheet. *Science*, 311(5763):986–990, 2006.

[34] Cecilie Rolstad, Jostein Amlien, Jon-Ove Hagen, and Bengt Lundén. Visible and near-infrared digital images for determination of ice velocities and surface elevation during a surge on osborne-breen, a tidewater glacier in svalbard. *Annals of Glaciology*, 24:255–261, 1997.

[35] R Rosenau, M Scheinert, and R Dietrich. A processing system to monitor greenland outlet glacier velocity variations at decadal and seasonal time scales utilizing the landsat imagery. *Remote Sensing of Environment*, 169:1–19, 2015.

[36] Vanessa Round, Silvan Leinss, Matthias Huss, Christoph Haemmig, and Irena Hajnsek. Surge dynamics and lake outbursts of kyagar glacier, karakoram. *The Cryosphere*, 11(2):723, 2017.

[37] Theodore A Scambos, Melanie J Dutkiewicz, Jeremy C Wilson, and Robert A Bindschadler. Application of image cross-correlation to the measurement of glacier velocity using satellite image data. *Remote sensing of environment*, 42(3):177–186, 1992.

[38] Dirk Scherler, Sébastien Leprince, and Manfred R Strecker. Glacier-surface velocities in alpine terrain from optical satellite imagery—accuracy improvement and quality assessment. *Remote Sensing of Environment*, 112(10):3806–3819, 2008.

[39] Dirk Scherler, Bodo Bookhagen, and Manfred R Strecker. Spatially variable response of himalayan glaciers to climate change affected by debris cover. *Nature geoscience*, 4(3):156–159, 2011.

[40] Robert A Schowengerdt. Measurement of the landsat thematic mapper modulation transfer function using an array of point sources. *Optical Engineering*, 27(4):334–343, 1988.

[41] Uttam Babu Shrestha, Shiva Gautam, and Kamaljit S Bawa. Widespread climate change in the himalayas and associated changes in local ecosystems. *PLoS One*, 7(5):e36741, 2012.

[42] James Storey, Michael Choate, and Kenton Lee. Landsat 8 operational land imager on-orbit geometric calibration and performance. *Remote sensing*, 6(11):11127–11152, 2014.

[43] James C Storey. Landsat 7 on-orbit modulation transfer function estimation. In *Proc. SPIE*, volume 4540, pages 50–61, 2001.

[44] USGS. Landsat geometry and data processing, June 2017. URL `https://landsat7.usgs.gov/geometry`.

[45] USGS. Slc-off gap-filled products gap-fill algorithm methodology, June 2017. URL `https://landsat.usgs.gov/sites/default/files/documents/L7SLCGapFilledMethod.pdf`.

[46] USGS. Special issue on landsat collections, June 2017. URL `https://landsat.usgs.gov/sites/default/files/documents/Special_Iss3_2016.pdf`.

[47] USGS. The worldwide reference system, June 2017. URL `https://landsat.gsfc.nasa.gov/the-worldwide-reference-system/`.

[48] USGS. Landsat 8 (l8) data users handbook, June 2017. URL `https://landsat.usgs.gov/landsat-8-l8-data-users-handbook-section-2`.

[49] Patrick Wagnon, Anurag Linda, Yves Arnaud, Rajesh Kumar, Parmanand Sharma, Christian Vincent, Jose George Pottakkal, Etienne Berthier, Alagappan Ramanathan, Syed Iqbal Hasnain, et al. Four years of mass balance on chhota shigri glacier, himachal pradesh, india, a new benchmark glacier in the western himalaya. *Journal of Glaciology*, 53(183):603–611, 2007.

[50] MGHY Winiger, M Gumpert, and H Yamout. Karakorum–hindukush–western himalaya: assessing high-altitude water resources. *Hydrological Processes*, 19(12):2329–2338, 2005.

[51] M Zemp, M Hoelzle, and W Haeberli. Six decades of glacier mass-balance observations: a review of the worldwide monitoring network. *Annals of Glaciology*, 50(50):101–111, 2009.

A

Appendix B

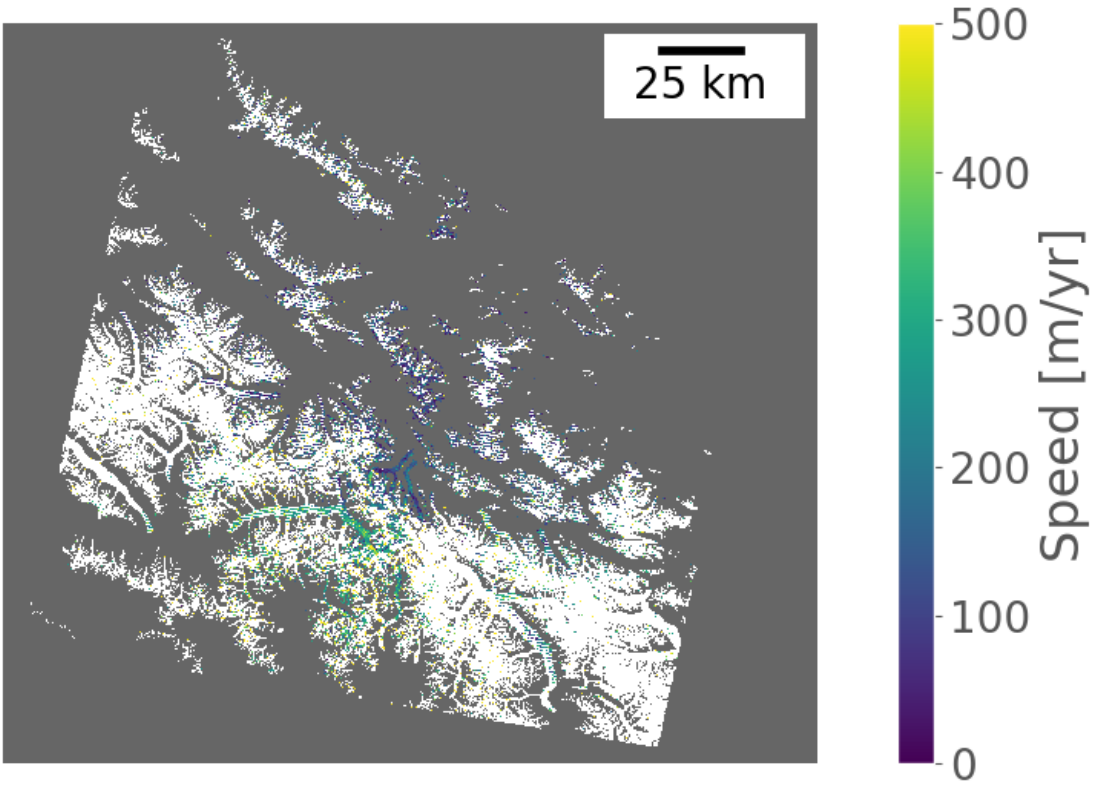


Figure A.1: Velocity fields from Landsat 7 SLC-off scenes in winter 2004 from images with 32 days temporal baselines. The data gaps are not white noise filled. The image is filtered with a SNR threshold of 3 and processed using a maximum velocity of 500 m/yr. The grey areas indicate the non-glacier masked areas.

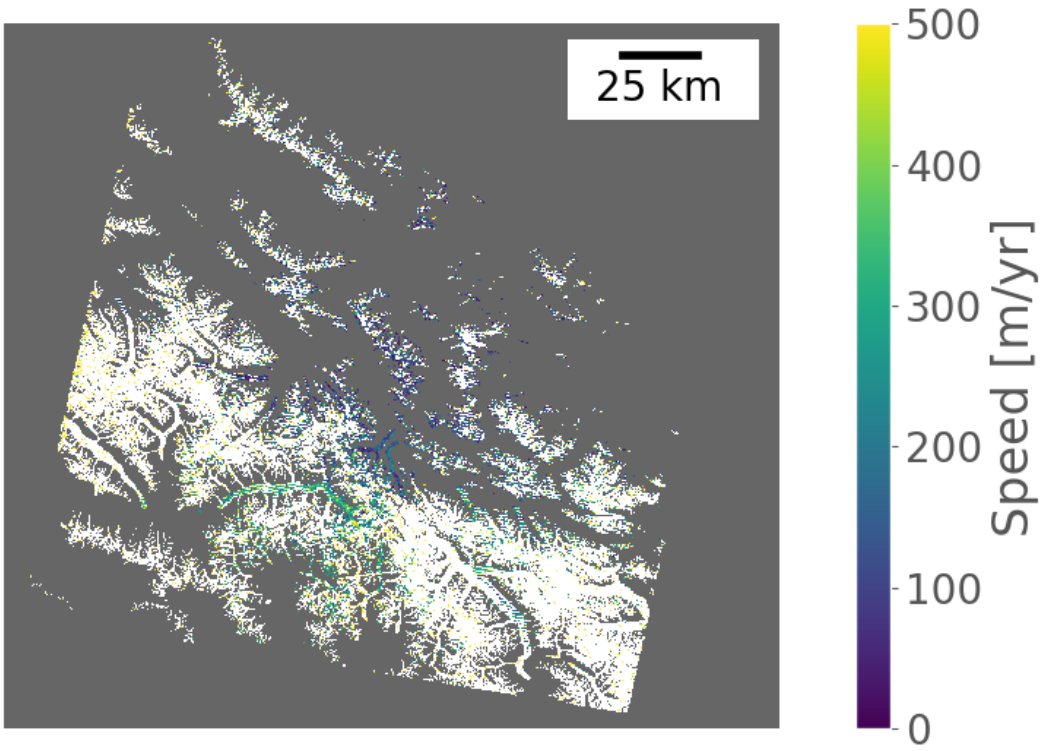


Figure A.2: Velocity fields from Landsat 7 SLC-off scenes in winter 2004 from images with 32 days temporal baselines. The data gaps are white noise filled. The image is filtered with a SNR threshold of 3 and processed using a maximum velocity of 500 m/yr. The grey areas indicate the non-glacier masked areas.

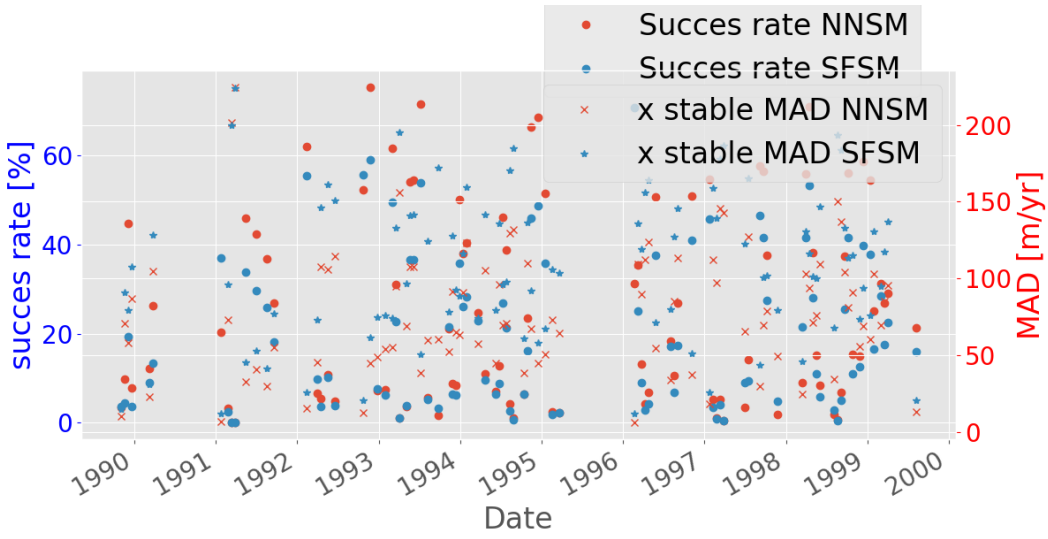


Figure A.3: Performance plot for time series of Landsat 5 for the Karakorum region (path/row 148/035). The left axis is the success rate percentage, the right axis the MAD of stable area in m/yr. On the x-axis the date of the first image used in the velocity pair. Comparison of the two tested methods.

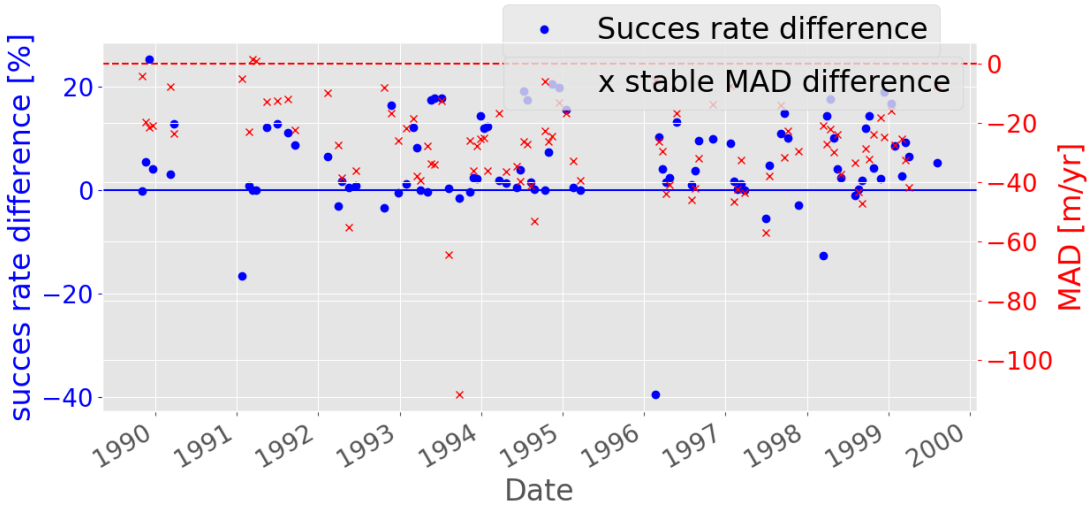


Figure A.4: Performance plot for time series of Landsat 5 for the Karakorum region (path/row 148/035). The left axis is the success rate percentage, the right axis the MAD of stable area in m/yr. On the x-axis the date of the first image used in the velocity pair. Plotted is the difference of performance between the two tested methods.