TU Delft

Delft University of Technology

Assessing Geomorphologic Processes with Permanent Laser Scanning

A Case Study on the Dutch Coast

Kuschnerus, M.

DOI
10.4233/uuid:31b3d8f8-1c0e-4a02-8089-d18034fa8850

Publication date
2024

Document Version
Final published version

Citation (APA)
Kuschnerus, M. (2024). *Assessing Geomorphologic Processes with Permanent Laser Scanning: A Case Study on the Dutch Coast*. [Dissertation (TU Delft), Delft University of Technology]. https://doi.org/10.4233/uuid:31b3d8f8-1c0e-4a02-8089-d18034fa8850

Important note
To cite this publication, please use the final published version (if applicable).
Please check the document version above.

Copyright
Other than for strictly personal use, it is not permitted to download, forward or distribute the text or part of it, without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license such as Creative Commons.

Takedown policy
Please contact us and provide details if you believe this document breaches copyrights.
We will remove access to the work immediately and investigate your claim.

This work is downloaded from Delft University of Technology.
For technical reasons the number of authors shown on this cover page is limited to a maximum of 10.

Assessing Geomorphologic Processes with Permanent Laser Scanning

- A Case Study on the Dutch Coast

ASSESSING GEOMORPHOLOGIC PROCESSES WITH PERMANENT LASER SCANNING

- A CASE STUDY ON THE DUTCH COAST

Dissertation

for the purpose of obtaining the degree of doctor
at Delft University of Technology
by the authority of the Rector Magnificus, prof. dr. ir. T.H.J.J. van der Hagen,
chair of the Board for Doctorates
to be defended publicly on Monday, 7 October 2024 at 10:00 o'clock

by

Mieke KUSCHNERUS

Master of Science Mathematik, Technische Universität Berlin, Germany
born in Gehrden, Germany.

This dissertation has been approved by the promotors.

Composition of the doctoral committee:

Rector Magnificus,	chairperson
Prof. dr. ir. R.F. Hanssen,	Delft University of Technology, *promotor*
Dr. R.C. Lindenbergh,	Delft University of Technology, *promotor*
Dr. ir. S. de Vries,	Delft University of Technology, *promotor*
Independent members:	
Prof. dr. K. Anders	Technical University of Munich, Germany
Prof. dr. C. Harmening	Karlsruhe Institute of Technology, Germany
Prof. dr. ir. P.J.M. van Oosterom	Delft University of Technology
Prof. dr. K.M. Wijnberg	University of Twente
Prof. dr.ir. A.J.H.M. Reniers,	Delft University of Technology, reserve member

Keywords: permanent laser scanning; time series; change detection; clustering; multiple hypothesis testing

Printed by: Ridderprint, www.ridderprint.nl

Cover: permanently installed laser scanner overlooking Noordwijk beach

Copyright © 2024 by M. Kuschnerus

ISBN 978-94-6506-413-0

An electronic copy of this dissertation is available at
https://repository.tudelft.nl/.

This research is part of the research programme CoastScan with project number 16352 which was (partly) financed by the Dutch Research Council (NWO).

to Ricard

and our wonderful children

Contents

Summary ix

Samenvatting xi

Zusammenfassung xiii

1 Introduction 1
- 1.1 Motivation . . . 1
- 1.2 Background: Monitoring Coastal Dynamics . . . 1
- 1.3 Permanent Laser Scanning . . . 3
- 1.4 4D Change Analysis . . . 5
- 1.5 Limitations . . . 7
- 1.6 Objectives and Research Questions . . . 7
- 1.7 Outline . . . 9

2 Minimal Detectable Displacement at Different Temporal Scales 11
- 2.1 Introduction . . . 13
- 2.2 Related Work . . . 14
- 2.3 Data Acquisition and Properties . . . 16
  - 2.3.1 Instrument Specifications and Settings . . . 16
  - 2.3.2 Data Availability . . . 17
  - 2.3.3 Study Area and Test Areas . . . 18
- 2.4 Method . . . 20
  - 2.4.1 Pre-processing Workflow . . . 20
  - 2.4.2 Error influences . . . 22
  - 2.4.3 Error estimation . . . 23
  - 2.4.4 Hypothesis Testing for Time Series . . . 24
  - 2.4.5 Test Quality . . . 26
  - 2.4.6 Minimal Detectable Bias . . . 27
- 2.5 Results . . . 28
  - 2.5.1 Effects of Environmental Conditions on Estimation Quality . . . 28
  - 2.5.2 Differentiating Dynamic Processes on the Beach . . . 31
- 2.6 Discussion . . . 35
  - 2.6.1 Additional effects on estimation quality . . . 35
  - 2.6.2 Potential Improvements of Set-up and Processing . . . 37
  - 2.6.3 Comparison to Existing Methods . . . 38
  - 2.6.4 Assumptions on Uncertainty . . . 40
- 2.7 Conclusion and Recommendations . . . 40

3 Coastal Change Patterns from Time Series Clustering **43**
3.1 Introduction . . . 45
3.2 The permanent laser scan data set . . . 48
3.3 Methods . . . 51
3.3.1 Time Series Extraction . . . 51
3.3.2 Distance Metrics . . . 51
3.3.3 Clustering Methods . . . 52
3.3.4 Evaluation Criteria . . . 56
3.4 Results . . . 59
3.4.1 Clustering . . . 59
3.4.2 Identification of Change Processes . . . 63
3.5 Discussion . . . 64
3.5.1 Distance Measures . . . 64
3.5.2 Clustering Methods . . . 65
3.5.3 Derivation of Change Processes . . . 66
3.6 Conclusions . . . 68

4 Identifying Topographic Changes at the Beach **69**
4.1 Introduction . . . 71
4.2 Field Site and Data Set . . . 73
4.2.1 Instrument Settings . . . 74
4.3 Generation of Inventory of Trends . . . 76
4.3.1 Corrected Point Clouds and Filtering . . . 76
4.3.2 Digital Elevation Models and Elevation Time Series . . . 78
4.3.3 4D Inventory of Trends of Surface Elevations . . . 79
4.3.4 Parameter extraction . . . 81
4.4 Results . . . 81
4.4.1 Inventory of trends . . . 82
4.4.2 Comparison of trend inventory with JarKus transects . . . 83
4.4.3 Categorisation of different topographic change processes . . . 83
4.4.4 Comparison of natural vs. anthropogenic processes . . . 88
4.5 Discussion . . . 90
4.6 Conclusions . . . 92

5 Conclusions and Recommendations **95**
5.1 Conclusions . . . 95
5.2 Main Contributions . . . 97
5.3 Recommendations for Future Work . . . 99

Bibliography **112**

Acknowledgements **113**

About the author **115**

SUMMARY

Coastal areas are sensitive ecosystems, which are important as natural habitat, recreational areas and for protection of the hinterland. Coastal monitoring is essential in the analysis and prediction of coastal development. Coastal monitoring has become more urgent due to climate change and its effects on sea level and frequency of storm events. Permanent laser scanning (PLS) provides a tool to acquire 3-dimensional point clouds of an area nearly continuously over extended periods of time. PLS bridges the gap between incidental, in-situ measurement campaigns with high spatial detail, and frequent monitoring via satellite data at lower spatial resolution. Methods to process the complex high-resolution permanent laser scanning data are needed to find and analyse the effects of geomorphological processes over extended periods of time and with high spatial detail. This dissertation deals with the development of methods for spatio-temporal data mining of large 4D data sets from permanent laser scanning for the application of coastal monitoring on sandy beaches.

Two data sets of hourly 3-dimensional point clouds, acquired over periods of six months and three years at two different locations on the Dutch coast were analysed, to identify and assess geomorphological dynamics at the sediment surface. Each data set contains up to 20 000 epochs capturing the dynamics of the sandy beaches in Kijkduin and Noordwijk.

In this thesis two methods are developed: the application of multiple hypothesis testing for the estimation of minimal detectable bias and the generation of a so-called inventory of trends, and the application of clustering algorithms for grouping elevation time series. The first method using multiple hypothesis testing provides a means to define the minimal detectable bias for an expected model behaviour of time series from permanent laser scanning. This method provides a new way to detect small but persistent and statistically significant changes in longer time series derived from 3-dimensional point clouds. Using multiple hypothesis testing allows to identify linear changes with slopes of 0.032 m/day and sudden changes in elevation of 0.031 m with a given discriminatory power of 80% and significance level of 5% in 24-hour time series.

In an additional step, multiple hypothesis testing is used to reduce the complex permanent laser scanning data set to an inventory of trends, which consists of linear pieces of time series, matching the predefined statistical models and corresponding parameters. This method is applied to find and analyse times and areas where specific processes such as storms, aeolian sand transport or bulldozer works occur. The inventory of trends is particularly effective for the detection of aeolian sand transport, which has been difficult to identify using other coastal observations because it causes small, gradual deformations at the sediment surface.

The second method uses clustering algorithms to identify areas which are subject to similar change patterns. These change patterns are then easily associated with underly-

ing physical and anthropogenic processes, mostly tidal induced changes and bulldozer works.

In summary, the developed methods allow to effectively detect deformations on sandy beaches and establish their origins, such as storms, tides, anthropogenic activities or aeolian sand transport, with a resolution and detail that has not been achieved until now. These results allow further analysis and interpretation of geomorphological coastal processes. For instance, the analysis of bulldozer works in our study area leads to the conclusion, that not only buildings themselves, but also the associated human interventions on the sandy beach around each building have a significant impact on coastal morphology and possibly lead to increased erosion.

SAMENVATTING

Kustgebieden zijn kwetsbare ecosystemen. Met name Nederlandse kusten zijn belangrijk als natuurgebied, als recreatiegebied en voor de bescherming van het achterland. Regelmatig monitoren van de kust is essentieel voor de analyse en de voorspelling van kustontwikkeling. Door klimaatverandering en de effecten daarvan op het zeeniveau en de frequentie van stormen wordt kustmonitoring steeds urgenter. Permanent laser scannen (PLS) is een recent hulpmiddel om gedurende langere perioden en bijna continu 3-dimensionale puntenwolken van een gebied te verkrijgen. PLS overbrugt daarmee de kloof tussen incidentele, in-situ meetcampagnes met een hoge ruimtelijke resolutie en frequent waarnemen vanuit satellieten met een lagere ruimtelijke resolutie. Maar er zijn wel methodes nodig om deze complexe permanente laserscandata met haar hoge resolutie en data volume te verwerken. Doel is met name het vinden en analyseren van de effecten van geomorfologische processen over langere perioden en met veel spatieel detail. Dit proefschrift behandelt de ontwikkeling van methoden voor spatio-temporele data mining van grote 4D permanente laser scan datasets voor kustmonitoring op zandstranden.

Uitgangspunt voor dit werk vormen twee datasets met 3D puntenwolken, die met een tussenpoos van steeds een uur werden verkregen gedurende zes maanden en drie jaar op twee verschillende locaties, Kijkduin en Noordwijk, aan de Nederlandse kust. Deze datasets, met in een geval 20.000 epochen, werden geanalyseerd om de morpho-dynamica van het strand in kaart te brengen.

Voor dit doel worden in dit proefschrift twee methodes ontwikkeld: ten eerste wordt statistisch testen ingezet voor de schatting van een minimaal detecteerbare afwijking en het genereren van een zogenaamde inventory of trends. Ten tweede worden zogenaamde cluster algoritmes gebruikt om op elkaar lijkende tijdreeksen van hoogte data te groeperen, in groepen van interpreteerbaar gedrag. De eerste methode biedt een manier om de minimaal detecteerbare afwijking te bepalen voor tijdreeksen uit permanente laserscanning met een vooraf gegeven gedrag. Deze methode biedt een nieuwe manier om kleine, maar langdurige en significant relevante veranderingen te detecteren in langere hoogte tijdreeksen verkregen uit een reeks van 3D puntenwolken. Met behulp van deze methode kunnen lineaire veranderingen met hellingen vanaf 0,032 m/dag en plotselinge veranderingen in hoogte vanaf 0,031 m worden geïdentificeerd met een gegeven onderscheidend vermogen van 80% en een significantieniveau van 5% in tijdreeksen van een dag bestaande uit 24 achtereenvolgende 3D puntenwolken.

Na ontwikkeling van deze methode van meervoudige hypothesetests wordt deze vervolgens gebruikt om de complexe permanente laserscan dataset te reduceren tot een inventory of trends. Deze inventaris bestaat, voor verschillende locaties op het strand, steeds uit lineaire stukken tijdreeks, die overeenkomen met de vooraf gedefinieerde statistische modellen en bijbehorende parameters. Deze methode wordt toegepast om tij-

den en gebieden te vinden, waar specifieke processen zoals stormen, wind gedreven zandtransport of bulldozerwerkzaamheden, starten, plaatsvinden en ook weer ophouden. Deze inventarisatie is vooral effectief voor de detectie van wind gedreven zandtransport, dat moeilijk te identificeren is met andere waarnemingsmethodes, omdat matige wind slechts kleine en geleidelijke verandering veroorzaakt. De tweede methode gebruikt cluster algoritmes om locaties bij elkaar te zoeken, die op een vergelijkbare manier veranderen. Deze gegroepeerde veranderingen kunnen vervolgens gemakkelijk in verband worden gebracht met de onderliggende fysische en antropogene processen, meestal veroorzaakt door het getij of door bulldozers.

Samenvattend kunnen we stellen dat de ontwikkelde methoden het mogelijk maken om veranderingen op zandstranden effectief te detecteren en bovendien de oorsprong van de veranderingen vast te stellen, zoals stormen, getij, menselijke activiteit of wind gedreven zandtransport. Bovendien maakt de combinatie van deze methodes met de gegeven dataset het mogelijk dit te doen met een resolutie en een mate van detail die tot nu toe nog niet bereikt waren. Deze resultaten maken de weg vrij voor verdere analyse en interpretatie van morpho-dynamische kustprocessen. De analyse van bulldozerwerkzaamheden in ons studiegebied leidt bijvoorbeeld tot de conclusie dat niet alleen de bebouwing zelf, maar ook het bijbehorend menselijke ingrepen op het strand rond elk gebouw een belangrijke invloed heeft op de kust en wellicht lokaal leidt tot meer erosie.

ZUSAMMENFASSUNG

Küstengebiete sind empfindliche Ökosysteme, die als natürlicher Lebensraum, als Erholungsgebiete und zum Schutz des Hinterlandes wichtig sind. Das regelmäßige Monitoring der Küstengebiete ist für die Analyse und Vorhersage künftiger Veränderungen von entscheidender Bedeutung. Infolge des Klimawandels und seiner Auswirkungen auf den steigenden Meeresspiegel und die Häufigkeit von Sturmereignissen, gewinnt das Monitoring der Küstengebiete zusätzlich an Bedeutung. Das permanente Laserscanning (PLS) ermöglicht die nahezu kontinuierliche Erfassung dreidimensionaler Punktwolken eines Gebietes über längere Zeiträume hinweg. Das PLS schließt die Lücke zwischen anlassbezogenen In-situ-Messkampagnen mit hoher räumlicher Detailgenauigkeit und der regelmäßigen Satellitenbeobachtung mit geringerer räumlicher Auflösung. Für die Verarbeitung der umfangreichen hochauflösenden permanenten Laserscanning-Daten werden Methoden benötigt, mit denen die Auswirkungen geomorphologischer Prozesse über längere Zeiträume und mit hoher räumlicher Detailgenauigkeit ermittelt und analysiert werden können. Diese Dissertation befasst sich mit der Entwicklung von Methoden zum raum-zeitlichen Data Mining umfangreicher 4D-Datensätzen aus dem permanenten Laserscanning für die Anwendung des Küstenmonitorings an Sandstränden.

An zwei verschiedenen Standorten an der niederländischen Küste wurden zwei Datensätze von stündlichen dreidimensionalen Punktwolken analysiert, die über einen Zeitraum von sechs Monaten bzw. drei Jahren erfasst wurden, um die geomorphologische Dynamik an der Sedimentoberfläche zu ermitteln und zu beurteilen. Jeder Datensatz enthält bis zu 20.000 Epochen, die die Dynamik der Sandstrände in Kijkduin und Noordwijk dokumentieren.

In der vorliegenden Arbeit werden zwei Methoden entwickelt: Die Anwendung multipler Hypothesentests für die Schätzung der minimal detektierbaren Abweichung und die Erstellung eines so genannten Inventory of trends, sowie die Anwendung von Clustering-Algorithmen zur Gruppierung von Zeitreihen der Höhe. Mit der ersten Methode, die multiple Hypothesentests verwendet, lässt sich die minimal detektierbare Abweichung für ein erwartetes Modellverhalten von Zeitreihen aus permanentem Laserscanning definieren. Diese Methode bietet eine neue Möglichkeit, minimale, aber anhaltende und statistisch signifikante Änderungen über längere Zeiträume aus dreidimensionalen Punktwolken zu detektieren. Mithilfe von multiplen Hypothesentests können lineare Veränderungen mit Steigungen von 0,032 m/Tag und plötzliche Höhenänderungen von 0,031 m mit einer gegebenen Trennschärfe von 80% und einem Signifikanzniveau von 5% in 24-Stunden-Zeitreihen erkannt werden.

In einem zusätzlichen Schritt werden multiple Hypothesentests eingesetzt, um den komplexen Datensatz des permanenten Laserscannings auf ein Inventory of trends zu reduzieren, das aus linearen Zeitreihen besteht, die den vordefinierten statistischen Modellen und zugehörigen Parametern entsprechen. Mit dieser Methode lassen sich Zeiten

und Zonen ausfindig machen und analysieren, in denen bestimmte Prozesse wie Stürme, äolischer Sandtransport oder Baggerarbeiten auftreten. Das Inventory of trends ist besonders effektiv für die Erkennung von äolischem Sandtransport, der mit anderen Küstenbeobachtungen nur schwer zu identifizieren ist, da dieser kleine, langsame Deformationen an der Sedimentoberfläche verursacht. Die zweite Methode verwendet Clustering-Algorithmen, um Gebiete zu identifizieren, die ähnlichen Veränderungsmustern ausgesetzt sind. Diese Veränderungsmuster lassen sich anschließend mit den zugrunde liegenden physikalischen und anthropogenen Prozessen in Verbindung bringen, vor allem mit gezeitenbedingten Veränderungen und Baggerarbeiten.

Die entwickelten Methoden ermöglichen es, Deformationen an Sandstränden effektiv zu erkennen und ihre Ursachen wie Stürme, Gezeiten, anthropogene Aktivitäten oder äolischen Sandtransport mit einer bisher nicht erreichten Auflösung und Detailgenauigkeit zu ermitteln. Diese Ergebnisse ermöglichen weitere Analysen und Interpretationen der geomorphologischen Küstenprozesse. Die Analyse der Baggerarbeiten in unserem Untersuchungsgebiet führt beispielsweise zu der Schlussfolgerung, dass nicht nur die Gebäude selbst, sondern auch die damit verbundenen anthropogenen Eingriffe in den Sandstrand um jedes Gebäude herum einen erheblichen Einfluss auf die Küstenmorphologie haben und möglicherweise zu verstärkter Erosion führen.

1

INTRODUCTION

1.1. MOTIVATION

Monitoring and analysing geomorphological processes in coastal areas is an integral part of research strategies to support coastal management. When monitoring at high spatial resolution is applied, it is typically done at low temporal resolution (e.g. yearly). PLS allows to acquire frequent (hourly) and high spatial resolution observations in large 4D data sets. In this work, methods for the exploitation and utilisation of these 4D data sets are developed and applied for the analysis of dynamic coastal processes.

1.2. BACKGROUND: MONITORING COASTAL DYNAMICS

Recent developments associated with climate change have made coastal areas increasingly vulnerable to extreme weather events. At the same time, a large part of the worldwide population is living in close proximity to coastal areas (A. Luijendijk *et al.*, 2018). Especially for a country like the Netherlands, with a large proportion of land immediately bordering coastal areas, sustainable coastal management is essential. In order to maintain the protective and the recreational function of coastal areas, it is very important to have a thorough understanding of processes that influence the evolution of coastal areas (Ranasinghe, 2016).

Natural geomorphologic processes like heavy storm events with strong winds and high tidal surges, currents and waves will likely increase in frequency due to climate change (Intergovernmental Panel on Climate Change (IPCC), 2014; Nicholls *et al.*, 2007). Coastal areas have a protective function during storm events and the immediate effects are researched and modelled extensively (see for example Callaghan *et al.*, 2013, Grabemann and Weisse, 2008). However, small-scale and slower processes like aeolian sand transport and erosion/deposition due to tidal forces are harder to monitor and analyse. They require consistent and frequent observation and accurate estimates of small elevation changes (centimetre level) over extended periods of time. These processes influence the geomorphologic long-term development of coastal areas. At the same time, these processes are potentially influenced by rising sea levels and interrupted and disturbed by more frequent storm events.

1

To protect coastal areas, 'building with nature' solutions like nourishments, which work with natural processes, rather than against them, are sometimes preferred over hard structures (see for example Brand *et al.*, 2022, Vriend *et al.*, 2015, Schipper *et al.*, 2016). One famous example is the 'sand motor', a mega-nourishment on the Dutch coast, which is replacing regular nourishments and making use of natural processes to transport sand to the required places (A. Luijendijk & Oudenhoven (Eds.), 2019). In view of this trend towards 'building with nature', the influence of buildings and human structures on (sandy) coasts has also been studied (see for example Bergen *et al.*, 2021 and Poppema *et al.*, 2021). In these studies, the focus is on the buildings themselves, and how their location and shape affects processes such as aeolian sand transport or dune growth. However, these studies rarely consider human activities that are concentrated around these buildings. Most buildings in coastal areas are frequently used and maintained and thus cause a large range of small-scale human interventions on the beach. Examples of these interventions include: creating embankments for a terrace, protecting a building with a sand dike before a storm and frequent clearing of access paths and wooden decks from sand (see Figure 1.1). The long term effect of the sum of many of these frequently repeated small-scale anthropogenic activities is underrepresented in coastal research. That is why the long term effects are unknown and difficult to predict.

Figure 1.1: Examples of natural and anthropogenic activities in the aeolian zone of Noordwijk beach (NL). A: Aeolian sand transport in front of embankment around beach bar. B: Access path covered with sand after a storm. C: Clearance of access path with a bulldozer.

Both natural and anthropogenic drivers need to be taken into account to analyse and model geomorphological coastal processes. Observation data of these highly dynamic areas is therefore needed. The Dutch coastal areas are regularly monitored with the help of remote sensing techniques. Such data is obtained in various ways, including satellite remote sensing (radar and optical), airborne laser scanning and close range inspection on the beach or from the water. Satellite observations using light detection and ranging

(LiDAR), spectral or radar sensors generally occur relatively frequent (daily/weekly) but with low spatial resolution (several 10s of meters). Airborne laser scanning and aerial photographs are taken yearly (Jarkus data set Rijkswaterstaat - Dutch Ministery of Infrastructure and Water Management, 2022) with a high spatial resolution (sub-meter). Inspections are carried out incidentally. However, often visual inspections are not exact and it is difficult to derive quantitative measurements for analysis. Photogrammetry from an instrument or a vehicle on the beach or a drone is used for incidental studies. Permanently installed cameras have been utilised for regular and high frequency observations, but with low spatial resolution and low level of detail (Davidson *et al.*, 2007; Holman & Stanley, 2007).

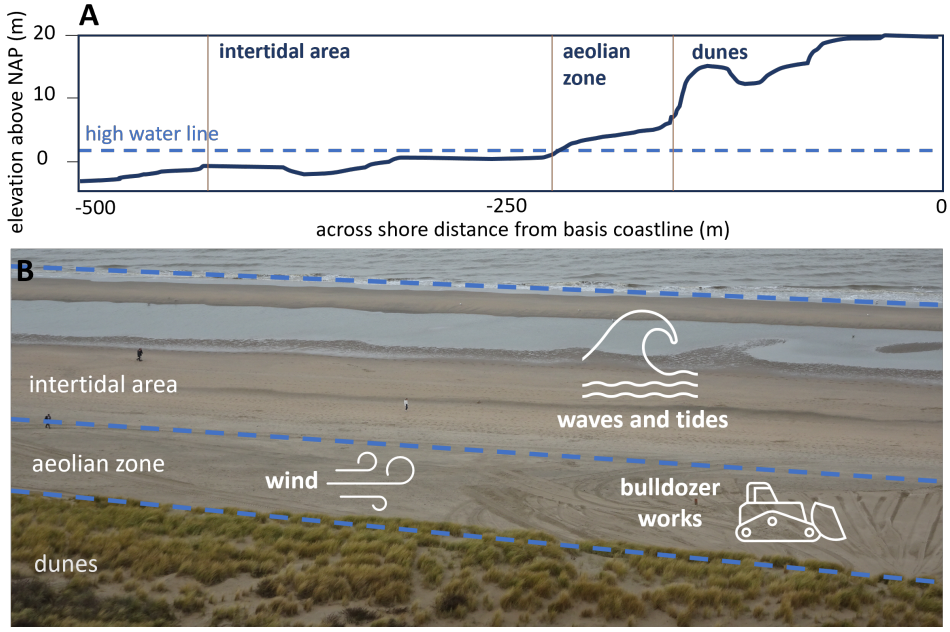

Figure 1.2: A: Beach profile with division into intertidal area, aeolian zone and dunes. B: Illustration of the main physical processes driving changes at the sediment surface in the different zones in sandy coastal areas.

1.3. PERMANENT LASER SCANNING

There is a gap between low frequency, high resolution observations from airborne remote sensing and incidental in-situ observations on one hand, and higher frequency, but low resolution observations from satellite remote sensing on the other hand. To bridge this data gap, permanent laser scanning (PLS) was introduced by Vos *et al.*, 2017: A terrestrial laser scanner (TLS) is mounted at a fixed position for an extended period of time, acquiring data of the same scene regularly and with high temporal frequency. At

each epoch a 3-dimensional point cloud is acquired using LiDAR, which represents the current shape of the observed area, as well as its reflectivity. LiDAR is a well established technique, which makes use of the travel time of an emitted laser pulse to estimate the distance of an object from the laser scanner. The emitted pulse is reflected and recorded by the laser scanner, which allows to derive the position of the reflecting point of the observed object in 3-dimensional coordinates. The collection of all these points forms a point cloud. For a schematic illustration of this technique see Figure 1.3 and for example Vosselman and Maas, 2010.

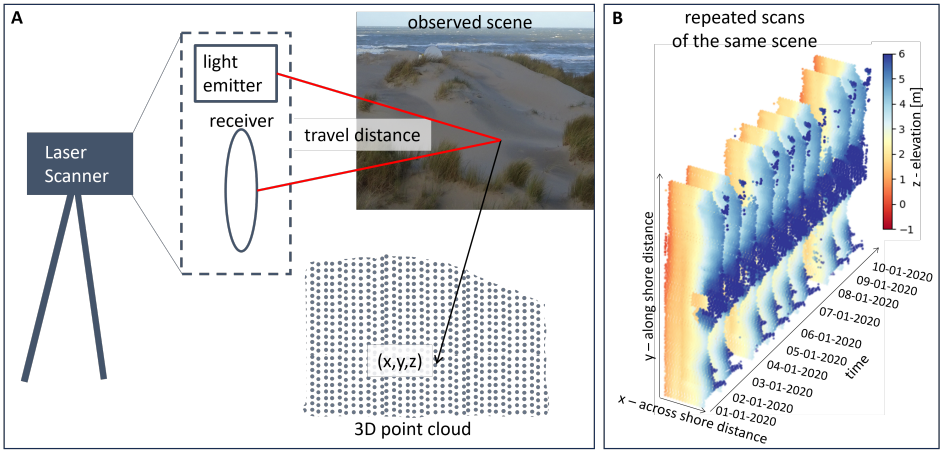

Figure 1.3: A: Illustration of PLS, which makes use of a permanently installed laser scanner using LiDAR to estimate 3D coordinates of an object. B: Resulting repeated point clouds of the same scene.

The 3D representations in each point cloud over time together form a 4-dimensional (3D + time) spatio-temporal data set. It contains the evolution of the observed area at a high level of detail (centimetre to meter scale) and with high temporal frequency (hourly to daily/weekly). Therefore dynamic processes like human maintenance activities in coastal areas, but also rock fall, landslide movement, snow melting on a glacier or vegetation development can be monitored closely.

PLS provides a tool to acquire a large 4D data set which potentially samples every change in the 3D geometry of the observed surface, as well as changes in the dielectric properties of the surface, potentially related to e.g., soil moisture. Recent examples of such data sets from PLS include

- The CoastScan data set (Vos *et al.*, 2017; Vos *et al.*, 2022) covering different parts of the Dutch and Belgian coast and facilitating analysis of dynamic coastal processes. This data set provides the input to the case studies and method development in this study.
- A data set collected in the boreal forest in Finland (Campos *et al.*, 2021) to observe

and study the dynamics of the forest canopy over several months.

- A data set collected at Hintereisferner in Austria (Voordendag *et al.*, 2023) to observe surface processes on a glacier over several years.

- A landslide monitoring site in Tirol, Austria. Monitored to derive surface changes, analyse potential risks and test long-term high-frequency monitoring over several months (Czerwonka-Schröder, 2023).

Each of these data sets serves a different application, but they share a common challenge: The large and unstructured data set consisting of a collection of point clouds is demanding to manage and deriving the relevant parameters to describe observed processes is not immediate. Dedicated methods to identify observed processes and extract parameters to describe them are needed in order to allow utilisation of the large amount of information contained in these data sets. Additionally, the size of the data sets poses computational challenges, if methods are not specifically adapted.

1.4. 4D CHANGE ANALYSIS

There are well established methods for the pairwise comparison of point clouds to support change detection in geomorphologic applications. One prominent example is the M3C2 algorithm developed by Lague *et al.*, 2013 to compare two point clouds of complex surfaces and determine statistically significant changes. When processing and comparing a large number of point clouds this is not an efficient method. The incorporation of the time dimension into the change analysis has been made more urgent by the advancement of techniques to acquire point clouds with high temporal frequencies and the various applications in long-term monitoring. These data sets contain a large amount of detailed information, but the extraction of a single process, or the analysis of a specific effect are no simple tasks and they do not rely simply on pairwise change detection. Therefore, there is a need to develop specific methods for 4D change analysis taking into account the large amount of estimates in four dimensions, as well as respective applications.

As the current state of the art, some methods used for comparison of several (few) epochs from terrestrial or airborne laser scanning are applied to extract some of these changes. Examples are the studies by Wheaton *et al.*, 2010 and Milan *et al.*, 2007 for comparison of differences using digital elevation models from point clouds. However, these methods need to be replicated for every epoch, which is inefficient. A 4D data set from PLS contains many epochs at high temporal sampling and smaller registration errors, due to the acquisition from a fixed position. Registration errors appear due to point clouds not being perfectly aligned, i.e. registered. On the other hand, the lack of stable reference surfaces in all three orientations and at different distances within the dynamic observed scenes, oftentimes hinders the efficient and accurate registration of all point clouds. Voordendag *et al.*, 2023 analysed the resulting error budget of a long-range PLS data set in a mountainous region including effects of environmental influences.

A study by Anders *et al.*, 2021 developed the so-called 4D-objects by change to extract processes and affected regions from a PLS data set. Since a 4D data set from permanent or frequent laser scanning is oftentimes large and challenging to analyse, they developed their method to segment such a data set in both space and time. Starting from so-called core points, their method finds pieces of time series that behave in similar ways and are in the spatial vicinity of the core point using dynamic time warping distance (Berndt & Clifford, 1994) to compare sections of time series. The segments derived through region growing starting from the core point are the 4D-objects by change and allow the analysis of drivers and influencing factors for observed processes as well as the analysis of patterns and typical change behaviour. In a further publication, Anders *et al.*, 2021 successfully demonstrate the use of their method for the detection of dynamic process on a sandy beach. Another application is the detection of erosion pattern of a snow covered alpine area (Anders *et al.*, 2022). However, the collection of the 4D-objects by change is not easily interpreted or analysed and favours most pronounced and spatially connected changes. Some long-term processes spread out over several locations might be missed. A method to cluster and bring order into 4D-objects by change is presented by Hulskemper *et al.*, 2022.

A study by Lindenbergh and Hanssen, 2003 uses testing theory to group time series of several epochs of elevation estimates from airborne laser scanning. They successfully detect and distinguish time series representing surface processes such as beach nourishments or marine erosion. A clear challenge was here the detection of start and end of a process and varying duration and time scales of surface activities, which did not always match the pre-defined length of the time series.

A different application was investigated by Weidner *et al.*, 2021, who developed a clustering approach to track trees in laser scanning data in order to monitor a landslide. The landslide movement was tracked with point clouds from laser scanning using an unsupervised decision tree algorithm to group relevant time series and subsequently estimate the observed displacement. Clustering of time series was also used by Puttonen *et al.*, 2019, who use nearest neighbours clustering to group time series of trees. The grouped time series are analysed to monitor and compare the circadian rhythms of different trees. Zoumpekas *et al.*, 2021 compare different supervised machine learning methods for the automatic detection of rock fall in a data set from frequent laser scanning covering several epochs. They find different methods are suitable for different cases, but establish two successful examples for the classification of rock fall events. They also conclude that each specific application needs reevaluation of the methods and parameters used and that general recommendations and strategies are not obvious from their case studies.

The main challenge for existing methods is to observe the generally unknown physical processes that are interacting which each other and often are superimposed in space and time. The nature of many of the observed change patterns on sandy beaches, but also in the snow, on top of a glacier, or on a landslide, could be described as a Langrangian process. In theory, the movement of each particle of sand (or snow, soil/rock) could be tracked. In practice however, these processes are observed as Eulerian changes, where

at a fixed location the evolution of height changes is observed and quantified. Consequently, the identification and separation of each dynamic process witch associated change pattern is challenging. The error estimation of each individual observation as well as the temporal evolution and correlation of observations in time and space add to the challenges. A specific opportunity arises from the long-term observation of small-scale and slowly developing processes.

1.5. LIMITATIONS

One of the original ideas at the start of the research project related to this study, was to use supervised machine learning (classification) or deep learning to detect and identify change pattern in the larger 4D PLS data sets. However, after some attempts and careful consideration, the focus of this study lies on the application and adaptation of other more conventional methods to be used for this purpose. The data sets provide a large amount of point clouds, but considering for example major storm events, three years of observations do not yield enough training data. With more understanding of the PLS data, the necessary (pre-)processing and the acquisition of more training data, the use of supervised machine learning might be a promising opportunity.

During the course of the research project, it became clear that many aspects of the acquisition and pre-processing of the data sets could potentially be improved. However, the analysis and application of the present data set was prioritised over the optimisation of data acquisition and set-up or registration method. An in depth error analysis including all correlation effects is also not part of this study and would require more additional data from other sensors.

1.6. OBJECTIVES AND RESEARCH QUESTIONS

The need for an efficient method to extract relevant parts of a 4D data set from PLS in order to provide information and interpretation on the observed processes is leading the research of this study.

How can geomorphologic processes be identified and quantified in a 4D spatio-temporal data set from permanent laser scanning?

This main research question is answered for an application in monitoring a typical part of the Dutch coast. At two different locations in the Netherlands a 4D spatio-temporal data set is acquired with PLS in order to study dynamic processes, support coastal management, and analyse the effects of small-scale, but frequent anthropogenic activities. The permanently installed laser scanner on the coast of Noordwijk, The Netherlands, is shown in Figure 1.4 together with a point cloud of the sandy beach and dune area.

To develop an exhaustive method for the extraction of geomorphologic processes from a 4D spatio-temporal data set for the exemplary case of the Dutch coast, the following sub-questions were addressed:

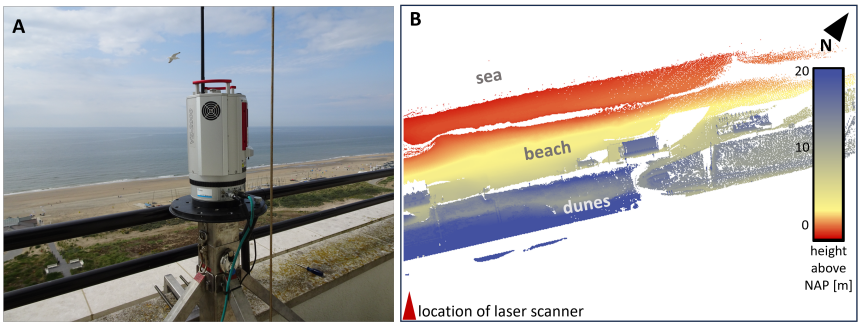

Figure 1.4: A: Permanently installed laser scanner overlooking the beach-dune system in Noordwijk, The Netherlands. B: Point cloud of the beach-dune system visible in A. The sandy beach appears in red and yellow and the dunes in blue.

(Q1) *What influences the uncertainty of height estimates from permanent laser scanning on a sandy beach?* (**Chapter 2**)

Even though the terrestrial laser scanner is recording range estimates, changes in height of the sediment surface are considered for this research. Influences on the uncertainty of height estimates from airborne and (mobile) terrestrial laser scanning have been analysed in detail in previous works (see for example Bitenc *et al.*, 2011). However, the specific challenges and opportunities of the permanently installed laser scanner set-up are still subject of ongoing research. The main influences on uncertainty of height estimates from PLS data were analysed and contributions to estimation errors when observing a dynamic scene on a sandy beach were quantified.

(Q2) *Which change patterns at the coastal sediment surface can be identified with permanent laser scanning?* (**Chapter 2 and 4**)

PLS allows to observe the sediment surface of a coastal area with high temporal and spatial resolution. This provides lots of opportunities to monitor coastal processes, such as aeolian sand transport, tidal erosion or human interference for example with bulldozer works. It is then investigated to which extent and under which conditions dynamic coastal processes can be monitored and quantified with the use of PLS. Here the focus lies specifically on the identification of the underlying driving forces (aeolian, tidal or anthropogenic) of slowly changing dynamics, such as erosion or deposition of sand over extended periods of time.

(Q3) *How can elevation time series resulting from the same process be grouped together?* (**Chapter 3**)

The combined point clouds from PLS allow to extract time series representing elevation evolution at the sediment surface. How these time series can be grouped together according to their behaviour, regardless of their location and in an efficient manner is

subject of this research. The focus of this part of our research lies on the grouping according to change patterns and temporal evolution. This allows to form clusters of all time series that result from the same physical process.

(Q4) *How are anthropogenic activities detected with permanent laser scanning and what is their influence on coastal development and natural dynamics?* (**Chapter 4**)

Anthropogenic activities, such as bulldozer works to restore parts of the sandy beach after a storm, clear access paths and terraces, or maintain recreational function of the beach are happening frequently on most parts of the Dutch coast. The influence of these relatively small-scale activities has been difficult to study with yearly observations and incidental inspections. PLS allows to analyse the impact of the activities for an exemplary part of the Dutch coast.

1.7. OUTLINE

The methods used in this study are summarised in Figure 1.5.

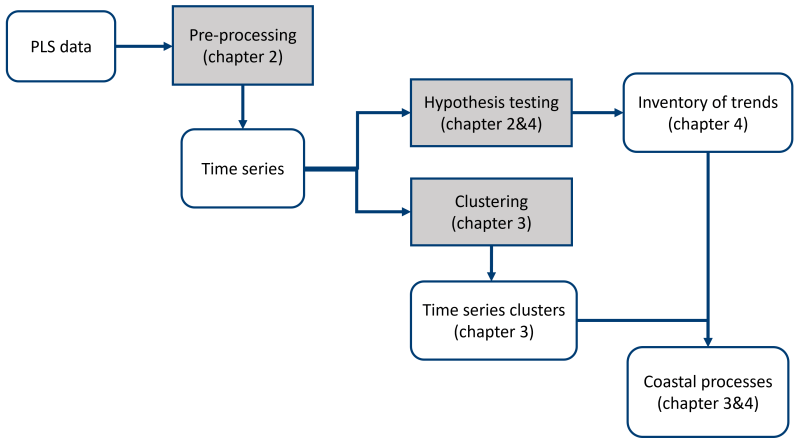

Figure 1.5: Overview of study methods and chapters. Methods are represented by grey boxes and data products are shown as white boxes with rounded corners.

Chapter 2 deals with influences on data quality for an example data set acquired in Noordwijk, The Netherlands, to answer question (Q1). Additionally, the minimal detectable displacement, measurable with a PLS system is introduced and multiple hypothesis testing (MHT) for time series from PLS data is discussed. It is shown how MHT can be used to categorise elevation time series. Further, it is addressed which processes can be identified using MHT and how temporal scale and frequency influence the results (research question (Q2)).

Clustering of elevation time series from another example PLS data set is discussed in **Chapter 3**. Time series of a fixed length are grouped together according to their temporal evolution. In this way, different change patterns of the sediment surface are found, as well as all locations where these occur. Here, question (Q3) is answered and different clustering methods on elevation time series are compared for an exemplary data set. Anthropogenic activities, such as clearing of paths and depositing piles of sand are detected as well.

Chapter 4 presents results on processing the entire three-year PLS data set acquired in Noordwijk, to extract change processes and differentiate between natural processes and human activities. The accumulated effect of human activities, as it can be observed in this case study is discussed as well. To answer the main question as well as questions (Q2) and (Q4) a collection of observed processes and their properties is generated as a so-called *inventory of trends*. Change processes identified as anthropogenic or aeolian are further analysed considering their cumulative and long term effect on the observed area.

Finally, in **Chapter 5** conclusions are drawn on the main research question and the application to the monitoring of the Dutch coast and recommendations for future research are provided.

2

Minimal Detectable Displacement at Different Temporal Scales in Permanent Laser Scanning Data on the Coast of Noordwijk, The Netherlands

Mieke Kuschnerus, Roderik Lindenbergh, Sander Vos, Ramon Hanssen

In the view of climate change, understanding and managing effects on coastal areas and adjacent cities is essential. Permanent Laser Scanning (PLS) is a successful technique to not only observe notably sandy coasts incidentally or once every year, but (nearly) continuously over extended periods of time. The collected point cloud observations form a 4D point cloud data set representing the evolution of the coast provide the opportunity to assess change processes at high level of detail. For an exemplary location in Noordwijk, The Netherlands, three years of hourly point clouds were acquired on a 1 km long section of a typical Dutch urban sandy beach. Often, the so-called level of detection is used to assess point cloud differences from two epochs. To explicitly incorporate the temporal dimension of the height estimates from the point cloud data set, we revisit statistical testing theory.

This chapter has been published in ISPRS Open Journal of Photogrammetry and Remote Sensing, **11**, 100055 (2024), Kuschnerus, Lindenbergh, Vos, and Hanssen, 2024.

We apply multiple hypothesis testing on elevation time series in order to identify different coastal processes, like aeolian sand transport or bulldozer works. We then estimate the minimal detectable bias for different alternative hypotheses, to quantify the minimal elevation change that can be estimated from the PLS observations over a certain period of time. Additionally, we analyse potential error sources and influences on the elevation estimations and provide orders of magnitudes and possible ways to deal with them. Finally we conclude that elevation time series from a long term PLS data set are a suitable input to identify aeolian sand transport with the help of multiple hypothesis testing. In our example case, slopes of 0.032 *m/day and sudden changes of* 0.031 *m can be identified with statistical power of* 80% *and with* 95% *significance in 24-hour time series on the upper beach. In the intertidal area the presented method allows to classify daily elevation time series over one month according to the dominating model (sudden change or linear trend) in either eroding or accreting behaviour.*

2.1. INTRODUCTION

In the view of climate change and the intensification of extreme weather events it is essential that coastal areas are monitored regularly with high accuracy. Permanent laser scanning (PLS) is an emerging measurement technique used to monitor natural areas including glaciers (Kellerer-Pirklbauer *et al.*, 2005), rockfall (Abellán *et al.*, 2010) and coasts (Vos *et al.*, 2017), as well as structures such as buildings, pipelines or mines (Mukupa *et al.*, 2017; Vezočnik *et al.*, 2009). PLS consists of a terrestrial laser scanner scanning frequently from a fixed position. With large amounts of point cloud data becoming increasingly manageable and improved instrumental set-ups, frequent terrestrial laser scanning (TLS), airborne laser scanning (ALS) as well as PLS are becoming well-established. PLS has the potential of detecting small scale changes in height or small deformations. The detectable changes reach centimetre levels (Anders *et al.*, 2019; Schröder & Nowacki, 2021; Vos *et al.*, 2022) and time scales of several days up to years are covered in different research projects (Schröder *et al.*, 2022; Voordendag *et al.*, 2021). However, at increased spatial and temporal resolution, environmental influences on the measurement system have a more pronounced effect on their performance (Kuschnerus, Schröder, *et al.*, 2021) and conventional methods for the determination of estimation quality, as developed for example for height estimations from ALS or TLS observations, do not always suffice.

A short analysis of error sources for height estimates from coastal PLS is presented by Vos *et al.*, 2020 and Kuschnerus, Schröder, *et al.*, 2021, who find that the strongest influence on the uncertainty in height estimates in a permanent coastal set-up during extreme weather conditions comes from precipitation and strong winds, which can both lead to data loss. Several methods are being developed to analyse geo-morphologically relevant processes in PLS data sets (see Anders *et al.*, 2021, Kuschnerus, Lindenbergh, and Vos, 2021, Campos *et al.*, 2021) without specifically considering quality of height or displacement estimates or minimal detectable changes in elevation or displacement. The *level of detection* as an indication of statistically significant surface change was developed for comparison between point clouds of two epochs by Brasington *et al.*, 2000 and Lane *et al.*, 2003. It is used to determine if differences between two point clouds in any direction are statistically significant for a fixed significance level (generally chosen as 95 %). The level of detection is used in combination with Kalman-filtering to detect significant height changes in 4D data sets by Winiwarter *et al.*, 2023 and for the comparison of differences between two epochs of rough surfaces with the M3C2 algorithm (Lague *et al.*, 2013). Methods that specifically consider multi-temporal point cloud comparison are still being developed and improved. Therefore the following research questions are posed to lead the research for this study:

- What is the minimal change in height on a sandy beach that can be estimated with a given confidence with our permanent laser scanning set-up?
- How do environmental conditions contribute to the uncertainty of height estimates from permanent laser scanning for the identification of change processes?
- Which change processes on a sandy beach can be observed with permanent laser scanning and at which temporal and spatial scales?

To answer these questions, we first introduce the properties of our specific PLS data set, followed by the processing steps leading to time series of digital elevation models (DEMs). Then we present the estimation of errors per grid cell of a DEM generated from each point cloud and the level of detection of height differences between scans. Further we adopt the multiple hypotheses testing methodology by Chang and Hanssen, 2016 for PLS data and use the model definitions for the estimation of the *minimal detectable bias* (MDB) in height estimation. The results cover geometric properties of the example data set, influences of environmental effects on the uncertainty of height estimation and a comparison of the concept of the level of detection with the MDB. Finally, we consider the detection of two geo-morphologic processes, aeolian sand transport and sudden changes caused by anthropogenic activities, and demonstrate the challenges of the presented method in the intertidal area.

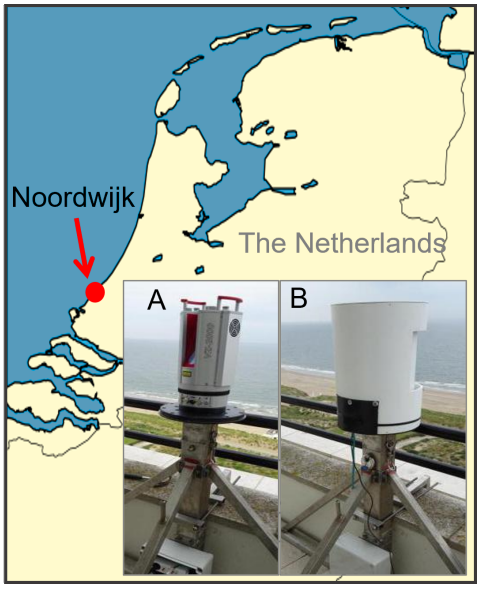

Figure 2.1: Location of the study site on the Dutch coast in Noordwijk, The Netherlands, and picture of the laser scanner mounted on a hotel balcony without (A) and with (B) protective cover.

2.2. RELATED WORK

Error sources in terrestrial laser scanning are summarised by Soudarissanane *et al.*, 2011. They emphasise the relevance of scanning geometry as influencing factor on height estimation quality and divide error sources into the following categories: scanner mechanism, atmospheric conditions and environment, object properties and scanning geometry. A review of commonly used performance estimation and error sources was presented by Muralikrishnan, 2021 with a focus on standardising the comparison of the quality of different instruments under test conditions.

Typically, accuracy of height estimations in digital elevation models (DEM) from laser scanning is assessed with the help of real-time kinematic positioning (RTK) GNSS measurements, as for example presented by Hladik and Alber, 2012 and Hodgson and Bresnahan, 2004. Bitenc *et al.*, 2011 use a theoretical error model in combination with overlapping LiDAR point clouds from mobile mapping to generate DEMs with corresponding error estimation. Rigorous theory on error models has been developed among others by Glennie, 2007 and by Lichti, 2007. These models require the acquisition of extensive self-calibration data sets under controlled laboratory conditions, which are not available in the case of many practical applications.

More recently Kerekes and Schwieger, 2020 developed an improved elementary error model (EEM) for height estimates from laser scanning with consideration of atmospheric effects on the measurements and therefore estimation accuracy. Winiwarter *et al.*, 2021 use error model theory to improve point cloud distance calculation based on a modified M3C2 algorithm as presented by Lague *et al.*, 2013, which was developed mainly for irregular rocky surfaces, as applied for example by Zoumpekas *et al.*, 2021 and for the comparison of two epochs. For more regular and flat surfaces Wheaton *et al.*, 2010 and Milan *et al.*, 2007 extend the theory of the level of detection to quantify uncertainty in DEMs of elevation differences (DoD) on river beds incorporating various error sources and error propagation. These techniques can be applied to a multi-epoch PLS data set as well, but the specific opportunities and challenges that arise from dealing with (nearly) continuous elevation time series have not been considered. A recent study by Voordendag *et al.*, 2023 discusses the five main influences on uncertainty of height estimates from PLS measurements of a glacier in the Alps. Their method is based on Soudarissanane, 2016 and derives the uncertainty in height estimation per DEM grid cell from single point measurements combined with registration errors. Williams *et al.*, 2018 consider a modified version of the M3C2 algorithm to detect volume change of rock falls. Williams *et al.*, 2018 conclude that higher temporal resolution can improve detection of instantaneous events, but simultaneously increases the accumulated estimation error for small-magnitude long-term processes.

Hypothesis testing is a well-established statistical technique. It was introduced by Baarda, 1968 for geodetic applications, and presented among others by Teunissen, 2006. Lindenbergh, 2010 use hypothesis testing to group and classify elevation times series from six consecutive point clouds. A rigorous approach on how to apply multiple hypotheses tests (MHT) on estimated kinematic InSAR time series was presented by Chang and Hanssen, 2016. They define a model data base to provide multiple alternative hypothesis and then use statistical testing, to identify the most likely kinematic model for each time series. The advantages and challenges of this method for the use on a large 4D data set from laser scanning of a very dynamic area, such as the coast, have not been discussed. We propose a simple way to use MHT for two basic models to identify geomorphologic processes on different time scales. The term *Minimal detectable bias* for MHT was first introduced by Baarda, 1968 and is discussed in detail by Imparato *et al.*, 2019. The MDB gives a measure of a minimum change in the estimated quantity of a time series that is needed in order to be identified with statistical significance.

For the comparison of two DEMs and calculating the statistical significance of the elevation difference per grid cell it is commonly assumed that both height estimates in the

respective grid cell are normally distributed. Adding the systematic registration error σ_{reg} and applying statistical testing using Gaussian statistics (based on Borradaile and Borradaile, 2003) results in the level of detection (LOD) at the 95% confidence interval as presented for example by Lague *et al.*, 2013. For the M3C2 algorithm a normal vector is determined from a neighbourhood around the considered point. This step can be skipped in the case of comparing two DEMs, since all height estimates of the DEMs are measured in z-direction. Thus, the necessity of the use of M3C2 or its advanced versions (Winiwarter *et al.*, 2021) is absent and the only remaining step is the estimation of the level of detection. The level of detection, LoD, between two grid cells in two DEMs with estimated variances σ_i^2 and number of points per grid cell n_i, $i \in \{1,2\}$ and registration error σ_{reg}, as defined by Lague *et al.*, 2013 is given by

$$\text{LoD} = \pm 1.96 \cdot \left(\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}} + \sigma_{reg} \right). \tag{2.1}$$

The estimated variances contain the terrain roughness as well as measurement precision. The level of detection is therefore a statistical confidence interval at a chosen significance level, here 95% for the difference between two independent surface elevations, with normally distributed measurement errors. This approach has also been applied to time series, where each height estimate per epoch was compared with the respective level of detection to a height estimate at the reference epoch, usually the first measurement (Winiwarter *et al.*, 2023).

2.3. DATA ACQUISITION AND PROPERTIES

2.3.1. INSTRUMENT SPECIFICATIONS AND SETTINGS

The point clouds are collected with a Riegl VZ-2000 laser scanner, which is permanently mounted on the balcony of Grand Hotel Huis ter Duin in Noordwijk, The Netherlands. The laser scanner is mounted on a metal frame at 55.757 m height above NAP (elevation above the Amsterdam Ordnance Datum) that is fixed to the balcony to maximise stability. The scanner is covered with a protective housing (see Figure 2.1) to shield it from rain, wind and dust.

range accuracy (at 150 m range) [m]	0.008
angular spacing [deg]	0.003
beam divergence [mrad]	0.3
wavelength [nm]	1550
inclination sensor	
measurement accuracy [deg]	0.008

Table 2.1: Specifications of Riegl VZ-2000 laser scanner according to documentation.

The specifications and settings of the laser scanner are summarised in Table 2.1. Each point cloud is generated by running a scan of nearly 180 degrees covering a part of the beach of just under 1 km every hour with angular spacing of 0.03 degree.

The system is set up with the same instrument and instrumental settings as a previous experiment at a different location, see Vos *et al.*, 2022.

2.3.2. DATA AVAILABILITY

area covered	250 000m^2
days scanned	954
interruptions (> 24h)	21
number of points per point cloud	$8.5 \cdot 10^6$
number of points (beach)	800 000
range (beach)	145 – 500 m
point density (beach)	1 – 40 pt/m^2
footprint size (beach)	0.015 – 0.27 m^2
incidence angle (beach)	70 – 90 deg

Table 2.2: Summary of data properties. The summary considers all hourly scans between 11 July 2019 and 21 June 2022.

The laser scanner generates one 3D point cloud per scan, made up of x, y, z-coordinates which are derived by the Riegl proprietary software out of observed range, horizontal and vertical angle data. The laser scanner also observes the intensity of the backscattered signal, per point. The internal inclination sensor records inclination values during each scan, with a frequency of 1 Hz (not matching the scanning frequency). These inclination angles are used for correction of tilts in the scanner (see section 2.4.1) and have a measurement accuracy of 0.008 degree.

The scanner operates 24 hours a day for the duration of three years. The numbers of available point clouds are visualised in Figure 2.2. A few gaps appear in the data collection. Some of them result from bad weather conditions, but most are due to technical failure and organisational problems. In May 2020 a decline in the data quality (i.e. increased presence of randomly located noise points) was observed and the scanner was finally switched off and sent for maintenance for a 34-day period at the end of June 2020 until end of July 2020. In December 2021 the scanner stopped working due to an unknown issue. Because of the holiday season it was not noticed until 18 days later in January 2022, when the scanner was restarted. The entire point cloud data set is published via 4TU Research Data (Vos *et al.*, 2023).

Additionally we collect data from a nearby weather station to separate environmental influences from other factors affecting the height estimation quality. We consider temperature, average wind speed per hour and precipitation as main influences on our height estimations. At the same time temperature, atmospheric pressure and humidity are provided to the instrument for internal range correction. These values are provided by meteoserver.nl and read from nearby local weather stations of the Royal Netherlands Meteorological Institute (KNMI) (Koninklijk Nederlands Meteorologisch Instituut, 2022) and updated hourly, to match the scanning frequency. For comparison and evaluation of results we use wind and temperature data from the KNMI weather station in Hoek van

Holland, on the coast at about 38 km distance from the laser scanner and to compare with visibility data we use measurements from a weather station in Schiphol, at about 25 km distance, but more inland.

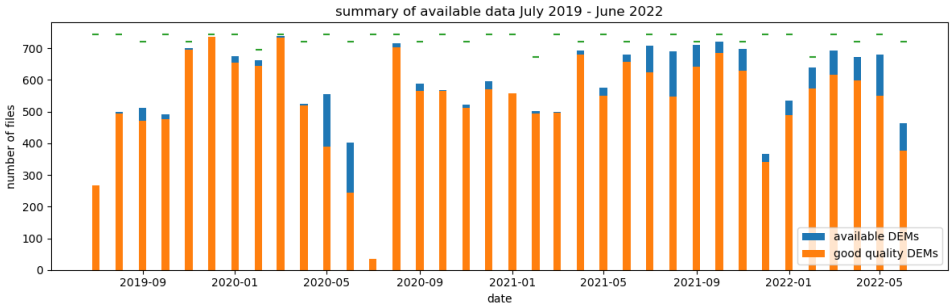

Figure 2.2: Available point clouds per month over the entire three-year period. The green dashes show the maximum number of files that would be possible per month. Two large gaps in the data collection are visible in June/July 2020 (maintenance) and in December 2021/January 2022, where the scanner stopped operating unnoticed during the holiday season.

2.3.3. STUDY AREA AND TEST AREAS

The observed area includes a sandy beach and dunes, covered with vegetation and is about 1 km long and 250 m wide. The beach is strongly influenced by the tides and varies in width between 80 m and 140 m under normal weather conditions. The area includes a helicopter landing platform (at 135 m range), which is used as a stable reference surface and a beach cafe at the dune foot on the sandy beach (at about 172 m range), which causes a large shadow area on the sandy beach, as shown in Figure 2.3.

For this study we focus on the sandy beach and disregard the dunes, parts of the hotel captured by the scanner and all points representing vegetation or other non ground objects (people, rubbish bins, etc.). A few reference surfaces (see Figure 2.3) from within the dune area are considered as well for the estimation of the registration error. The height above NAP of the reference surfaces has been confirmed with RTK-GNSS measurements and double checked with height estimates from the AHN (Actueel Hoogtebestand Nederland) measurement campaigns (GeoTiles.nl, 2021) using ALS. The range of the sandy beach varies within the point cloud between 150 and 500 m. On the beach we selected two exemplary test areas, as marked in Figure 2.3. Test area 1, at about 180 m range is representative of the dry part of the beach, where the tide does not reach under normal weather conditions. Because of the location right next to the beach cafe, it is subject to frequent bulldozer works and human activities. Test area 2, at about 290 m range, appears only a few times a day in the point clouds, as it is regularly submerged during high tide. Point spacing varies between 1 and 40 points per m^2 with non-overlapping, ellipse shaped footprints (short diameter between 0.04 and 0.08 m and long diameter between 0.11 and 0.8 m) with sizes ranging from 0.015 m^2 to 0.27 m^2. We assume that the height

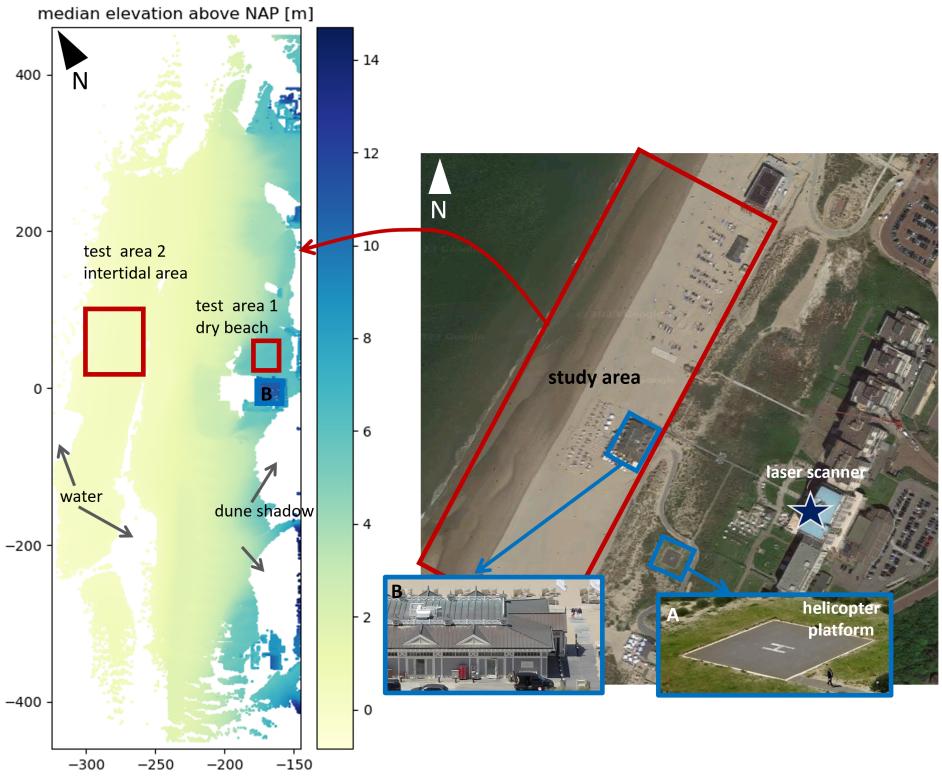

Figure 2.3: Top view of DEM of the area of interest at low tide on 14-04-2020 (left panel) with marked test areas on the dry beach and in the intertidal area. White areas in the DEM represent shadows of the dunes, buildings or flooded parts in the intertidal area, where the scanner is not recording any points. The x- and y-axis represent across-shore and along-shore distance in meters. Right panel: Overview of study site, located on the beach in Noordwijk (Google maps) with helicopter platform (A) and beach cafe (B) which serve as stable reference surfaces. The location of the laser scanner is marked with a star.

estimation per point represents the estimated height at the centre location of the foot-print. The incidence angle is rather unfavourable due to a surface slope of about 1 degree (on average) towards the sea and the position of the laser scanner. It ranges between 72 and 80 degree on the sandy part of the beach. A schematic of the side view of the set up and distances to the beach is shown in Figure 2.4 together with an illustration of the number of points and foot print size within a square meter at 145 m range.

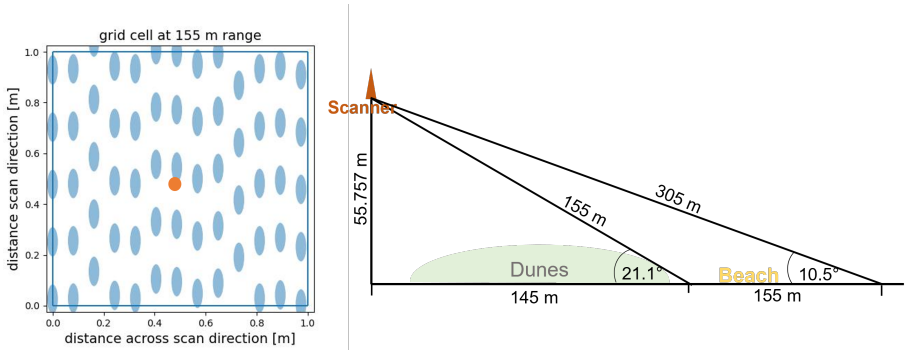

Figure 2.4: Left: Illustration of estimated footprint size and distribution within a square meter on the closest part of the beach (at about 155 m range). All height estimates within a square meter are averaged to estimate the height at the grid cell centre (marked in orange) to generate a DEM (see Section 2.4.1). Right: Schematic side view of the instrument set up and range to the beach at different horizontal distances.

2.4. METHOD

2.4.1. PRE-PROCESSING WORKFLOW

The workflow is shown as a schematic in Figure 2.5 and explained in the following paragraphs.

PREPARATION OF DATA

Each 3D point cloud is in a local coordinate system with the location of the laser scanner as its origin. We determined the height of the laser scanner when it was mounted with the help of a GNSS receiver. This constant elevation of 55.757 m above NAP is added to the z-coordinate, to process actual height above sea level instead of negative elevation with respect to the scanner's location.

In a first step, all points representing the hotel are removed from each point cloud. This leads to a reduction in size of each point cloud from about 5 Mio points to about 3.5 Mio relevant points.

CORRECTION OF SCANNER TILT

In a next step the 1 Hz inclination measurements from the laser scanner are averaged for each scan and the mean pitch and roll inclination are estimated. These values are then used to calculate a rotation matrix **R**:

$$\mathbf{R} = \begin{pmatrix} \cos(\varphi) & 0 & \sin(\varphi) \\ \sin(\varphi)\sin(\theta) & \cos(\theta) & -\cos(\varphi)\sin(\theta) \\ -\sin(\varphi)\cos(\theta) & \sin(\theta) & \cos(\varphi)\cos(\theta) \end{pmatrix} \tag{2.2}$$

with φ the pitch angle and θ the roll angle. The rotation matrix **R** is calculated for each point cloud separately and applied only, if the standard deviations of the pitch and roll

values during the entire acquisition of the respective point cloud are below a threshold. This is to ensure that point clouds acquired during a heavy storm are not corrected with a rotation based on erroneous inclination values. In these cases a mean pitch and roll value based on the other scans in that month is used for correction. A constant general tilt of the point cloud is removed in this way, as well as the main part of deviations of the laser scanner's position due to temperature changes (heat expansion of supporting materials) or strong winds (see Section 2.5.1).

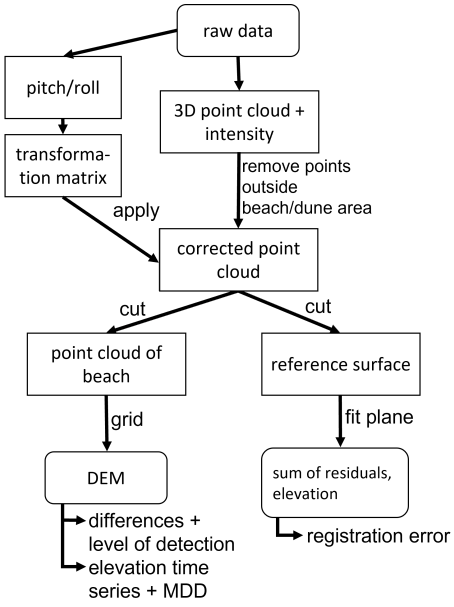

Figure 2.5: Flow chart of the pre-processing steps.

DEM GENERATION

For the following analysis of the systematic error affecting each point cloud and for a quality check, the reference surfaces are cut out using their x- and y-coordinates. We assume that the reference surfaces are flat with constant elevation and no tilt in x- or y-direction. Plane fitting provides the mean elevation of the respective reference surface as well as the squared sum of residuals between all points and the fitted plane. Both are used for a quality check: If the mean elevation of the plane deviates more than 0.1 m from the (GNSS verified) expected elevation or the squared sum of residuals is above 0.1 m^2, the respective point cloud is marked as 'bad quality' and not considered for further analysis. In Figure 2.2 these files are indicated as available (blue) but not included as 'good quality DEM' (orange).

For further analysis of elevation changes on the beach and in the test areas, all other parts of the point cloud are removed and subsequently a DEM is generated from the remaining points, covering only the sandy part of the beach. The DEM has a 1m × 1m

grid cell size and the mean elevation of all points is used as grid cell elevation estimate located at the centre of the grid cell (Figure 2.4, left). For each grid cell the accumulation of the estimated elevations over time provides the elevation time series at that location.

2.4.2. ERROR INFLUENCES

For further analysis we investigate the random error affecting each grid cell of the DEM, σ_g, and the systematic error affecting the entire point cloud, ε_{pc}. To summarise and quantify the most relevant error influences we estimate the order of magnitude for each of them following Soudarissanane, 2016 and Voordendag *et al.*, 2023, and distinguish the main influences: geometry, registration, atmosphere, instrument and surface properties.

GEOMETRY AND REGISTRATION

We assume that the measurement geometry does not change between scans, except for slight movement of the scanner, due to strong wind and movement of the entire hotel building due to concrete expansion. After correction of scanner tilt (see Section 2.4.1), the largest part of the geometric error is removed. A small error, due to the limited measurement accuracy of the inclination sensor will remain and we do not have the means to quantify any changes in yaw-angle. Since the start and stop angle of each scan are settings that are kept constant, we do not take into account any errors due to inaccuracies in yaw direction. The effect of concrete expansion can be estimated from temperature observations using the expansion coefficient (see for example Marshall, 1972) and is estimated to be below 1 cm. However, direct correlation with our measurements and correction of this error prove to be difficult. We do not take into account variations of incidence angle and footprint size, which both vary within the scene as well as over time with changes in the surface topography (see Table 2.2).

No additional registration step is applied. Therefore, small registration errors are still present when comparing subsequent point clouds. We observe the registration error on elevation measurements by analysing the height of a fitted plane through the stable reference surfaces. The small variations observed here (in the order of 1.5 cm) are the result of slight displacements of the scanner, scanner tilt and expansion of the building where the scanner is mounted on and potentially disturbances of the atmosphere. We cannot separate these effects and therefore summarise them in one error term ε_{pc} including all error sources affecting the entire point cloud systematically.

ATMOSPHERE

Atmospheric effects were found to influence the height estimates as reported by Kuschnerus, Schröder, *et al.*, 2021. Deviations in range and therefore in elevation can be caused by temperature gradients in the air and therefore differences in refraction index between the scanner location and right above the beach. The order of magnitude can be estimated empirically from elevation time series of the reference surfaces. Assuming the atmosphere does not vary within the area of interest, the atmospheric effects are part of the systematic error ε_{pc}. Low visibility due to fog prevents measurements all together

and leads to exclusion of the respective point cloud. For more explanation on the effect of atmospheric changes on height estimates from PLS see Voordendag *et al.*, 2023.

INSTRUMENT

Instrument errors can result from environmental effects: temperature fluctuations can influence the instrument. It is not known what the exact temperature under the protective cover and inside the instrument is at the time of data acquisition. Therefore, the temperature from a nearby weather station is possibly not representative for the temperature inside the laser scanner. Additionally, the laser scanner showed some erroneous behaviour in spring and early summer 2020. Point clouds acquired at this time contained more and more additional random points above and below the actual measured surface. These points were for a large part filtered out. The erroneous measurements could be part of the grid cell uncertainty σ_g. However, when filtering for mean elevation and summed residuals on the reference surfaces as explained in Section 2.3.2, the point clouds heavily affected by the malfunction, were excluded. Calibration of the instrument and long-term drifts in calibration parameters could potentially influence our measurements as well. However, two calibration reports at the beginning of the scan and after maintenance in summer 2020 did not show any significant changes.

SURFACE PROPERTIES

Surface properties in our test areas are relatively consistent, since we are interested in observing the sandy beach. They vary, however in soil moisture and surface roughness. Soil moisture content has an effect on reflectivity (Di Biase *et al.*, 2021; Jin *et al.*, 2021) and potentially affects range measurements. Surface roughness was further analysed by Di Biase *et al.*, 2022 where different grid cell size and variogramms of surface roughness were considered. All error terms affecting the grid cell are summarised as σ_g (equation 2.3).

2.4.3. ERROR ESTIMATION

We estimate the elevation in z-direction per grid cell as the mean elevation over all points within the grid cell. This corresponds to the simplified plane equation, where the estimated elevation $\hat{z}$ is just the mean of all observed values z_i from the vector of observations $\underline{\mathbf{z}}$ as in Bitenc *et al.*, 2011: $\hat{z} = \text{mean}(\underline{\mathbf{z}})$. The surface roughness and measurement uncertainty are then incorporated in the standard deviation of all points in the grid cell, which corresponds to the root mean square error (RSME) σ_g. We define σ_g by

$$\sigma_g = \sqrt{\frac{\underline{\mathbf{e}}_s^T \underline{\mathbf{e}}_s}{n_p - 1}}, \tag{2.3}$$

for the spatial residual vector $\underline{\mathbf{e}}_s$ of all elevation values in the grid cell and number of points $n_p - 1$, to ensure an unbiased estimation.

The systematic errors are dominated by the registration error plus temperature and atmospheric effects as explained above. It is estimated from the mean elevation of the reference surfaces and does not vary significantly with range, because the rotation part

is largely corrected and the remaining translations affect the entire point cloud independent of location. We assume that the observed standard deviation of the elevation of the fitted plane provides the systematic error $\varepsilon_{pc} = 1.5$ cm for the entire point cloud.

2

2.4.4. HYPOTHESIS TESTING FOR TIME SERIES

To allow classification of time series according to likely deformation models, we apply MHT based on the approach by Chang and Hanssen, 2016. More details and background can be found in the work of Teunissen, 2006 and some examples in Tiberius *et al.*, 2021. MHT allows to test for the null-hypothesis (no change) and several alternative hypotheses.

With the definitions as above, we consider the entire time series of one grid cell location, with elevation vector $\underline{\mathbf{y}} = (\hat{z}_1, \hat{z}_2 \ldots, \hat{z}_m)$ over time $\underline{\mathbf{t}}$ of length m and corresponding vectors of standard deviations $\underline{\sigma}_g$ and number of points per grid cell $\underline{\mathbf{n}}_p$.

We assume that the elevation vector $\underline{\mathbf{y}}$ can be estimated with the model $\underline{\mathbf{y}} = \mathbf{A} \cdot \underline{\mathbf{x}} + \underline{\mathbf{e}}_t$ for an unknown parameter vector $\underline{\mathbf{x}}$ of dimension n. The temporal residual vector $\underline{\hat{\mathbf{e}}}_t = \underline{\mathbf{y}} - \mathbf{A} \cdot \underline{\hat{\mathbf{x}}}$, is the estimated temporal residual vector not related to the spatial residuals in Section 2.4.3. We assume that the single elevation measurements are normally distributed as described above, and therefore the residual vector $\underline{\hat{\mathbf{e}}}_t$ is normally distributed with zero mean (see for example Tiberius *et al.*, 2021 for more details).

Now we define the null-hypothesis H_0 and one alternative hypothesis H_1 following Chang and Hanssen, 2016:

$$H_0 : \mathrm{E}(\underline{\mathbf{y}}) = A_0 \cdot \underline{\mathbf{x}} \tag{2.4}$$

$$D_y = Q_{yy} = \mathrm{diag}(\underline{\sigma}) \tag{2.5}$$

$$H_1 : \mathrm{E}(\underline{\mathbf{y}}) = A_0 \cdot \underline{\mathbf{x}} + C_1 \underline{\mathbf{b}}_1 \tag{2.6}$$

$$D_y = Q_{yy} \tag{2.7}$$

where $\mathbf{A}_0 \in \mathbb{R}^{m \times n}$ is the design matrix, i.e. representing the model of the null-hypothesis, $E(\cdot)$ the expectation operator and $\mathbf{D}_y$ is the dispersion which equals the covariance $\mathbf{Q}_{yy} \in \mathbb{R}^{m \times m}$ of the elevation vector $\underline{\mathbf{y}}$. For the alternative hypothesis H_1, the model is extended by a specification matrix $\mathbf{C}_1$ with $\underline{\mathbf{b}}_1$ the additional vector of unknown parameters. To either sustain or reject the null-hypothesis, considering the normal distribution of the residual vector $\underline{\hat{\mathbf{e}}}_t$, we use the χ^2-distribution, $\chi^2(q, \lambda)$, with non-centrality parameter λ and number of additional parameters of the alternative hypotheses q. For the null-hypothesis, the model is simply the mean elevation over the entire time series. In this case the model matrix $\mathbf{A}_0$ reduces to an m-dimensional vector, $n = 1$, yielding null-hypothesis H_0 and test value T_0

$$H_0 : \mathbf{A}_0 = \begin{pmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{pmatrix}, \quad x = x_0 \tag{2.8}$$

$$T_0 \sim \chi^2(1, 0). \tag{2.9}$$

As first alternative hypothesis we use a step function defined to test for sudden changes at time $\hat{t}$, with elevation $x_0 = \text{mean}(y_1 \dots y_{\hat{t}-1})$ for $t < \hat{t}$ and $x_1 = \text{mean}(y_{\hat{t}} \dots y_m)$ for $t \geq \hat{t}$, and dimensions $n = 2$ and additional parameter dimensions $q = 1$, yielding alternative hypothesis $H_{1,\hat{t}}$ and test value $T_{1,\hat{t}}$:

$$H_{1,\hat{t}} : \mathbf{A}_{1,\hat{t}} = \begin{pmatrix} 1 & 0 \\ \vdots & \vdots \\ 1 & 0 \\ 0 & 1 \\ \vdots & \vdots \\ 0 & 1 \end{pmatrix}, \quad x = \begin{pmatrix} x_0 \\ x_1 \end{pmatrix} \tag{2.10}$$

$$T_{1,\hat{t}} \sim \chi^2(1, \lambda). \tag{2.11}$$

With $\mathbf{C}_{1,\hat{t}} = (0, \cdots, 0, 1, \cdots, 1)^T$ and $b_{1,\hat{t}} = x_1 - x_0$ this is equivalent to our previous definition of the alternative hypothesis:

$$H_{1,\hat{t}} : \mathrm{E}(\underline{\mathbf{y}}) = \mathbf{A}_0 \cdot \underline{\mathbf{x}} + \mathbf{C}_{1,\hat{t}} b_{1,\hat{t}}. \tag{2.12}$$

The determination of $\hat{t}$ in practice is explained in more detail in Section 2.4.6.

And as second alternative hypothesis H_2 we use a linear trend and calculate test value T_2:

$$H_2 : \mathbf{A}_2 = \begin{pmatrix} 1 & t_1 \\ \vdots & \vdots \\ 1 & t_m \end{pmatrix}, \quad x = \begin{pmatrix} a_0 \\ a_1 \end{pmatrix} \tag{2.13}$$

$$T_2 \sim \chi^2(1, \lambda) \tag{2.14}$$

where a_0 is the intercept and a_1 the slope of the linear model and therefore $\mathbf{A}_2 \in \mathbb{R}^{m \times 2}$, $n = 2$ and $q = 1$. This is consistent with equation 2.6 when defining:

$$\mathbf{C}_2 = (t_1, \cdots, t_m)^T, \quad b_2 = a_1. \tag{2.15}$$

The covariance matrix is defined by the RMSE per grid cell (σ_g as defined in Section 2.4.3). We assume that there is no correlation between subsequent measurements and therefore obtain the diagonal matrix

$$\mathbf{Q}_{yy} = \text{diag}(\sigma_{g1}^2, \dots, \sigma_{gm}^2) = \underline{\sigma^2} \cdot \mathbf{I}_m. \tag{2.16}$$

Following Tiberius *et al.*, 2021 and Chang and Hanssen, 2016 the test value T_j, $j \in \{0, \dots, m+1\}$ for time series of length m for all alternative hypotheses, is then calculated as

$$T_j = \underline{\hat{\mathbf{e}}}_{j,t} \, \mathbf{Q}_{yy}^{-1} \, \underline{\hat{\mathbf{e}}}_{j,t} = \sum_{i=1}^{m} \frac{(y_i - \hat{y}_j)^2}{\sigma_i^2}, \tag{2.17}$$

with residual vector $\underline{\hat{\mathbf{e}}}_{j,t}$ for modelled time series $\underline{\hat{\mathbf{y}}}_j = \mathbf{A}_j \cdot \hat{x}$. The test value is then compared to the critical value k_α, for significance level α following the respective χ^2-distribution, with non-centrality parameter λ.

We test for every grid cell, if for a time series over a fixed amount of time (for example 24 hours) the null-hypothesis is not rejected. If it is rejected, we test in a next step, which of the alternative hypothesis is most likely i.e. has smallest value for T_j below the critical value. If T_j is above the critical value for all alternative hypotheses, no adequate model can be found. In this way sudden changes, as well as gradual, linear processes can be identified with a pre-defined level of confidence α. This method is suitable to be extended for periodic/seasonal changes or any other typical behaviour that one would expect on a sandy beach.

2.4.5. TEST QUALITY

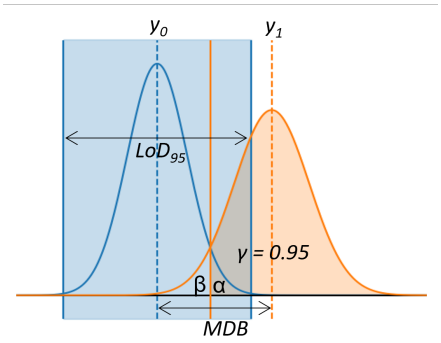

Figure 2.6: Comparison of level of detection with detectability power. The parameters α and β are indicated and they both equal 0.05 in this example. It can be seen that the level of detection with significance level 95% does not match the MDB with 95% detectability power.

The type I error describes the rejection of the null hypothesis, while it is true. The size of the type I error is the probability of this happening and is defined by significance level α. The size of the type II error is given by the probability of a missed detection of change (in height in our case), β. Its compliment is the detectability power $\gamma = 1 - \beta$, see Figure 2.6 for an illustration. The determination of the detectability power γ therefore depends on the calculation of the probability of a missed detection β, which is defined as the integral of the probability density function of the alternative hypothesis over the acceptance region of the null-hypothesis.

For the MHT, following the core idea of Baarda's B-method, also applied by Chang and Hanssen, 2016, the detectability power is fixed, for example at 80% and the significance level α can be defined depending on the number of parameters of the respective hypothesis, as well as the dimension of the time series. Then, the critical value T_j will be evaluated in relation to the significance level and the ratios will be compared instead of the critical values, since the significance level depends on the number of parameters of the alternative hypothesis. Here, we choose two simple models for the alternative hypothesis, which both have two parameters that need to be estimated. Therefore the critical values can be compared directly and the discriminatory power, and consequently the

type II error will be the same for all alternative hypothesis with the above definitions.

2.4.6. MINIMAL DETECTABLE BIAS

To answer our first research question, we use the concept of the *minimal detectable bias*, as introduced by Baarda, 1968, assuming we know which model(s) would best represent possible changes in elevation (alternative hypotheses) in a specific area and time period. In our case, we are looking for the minimal height of a step in a sudden change (step function model) and the minimal slope (height per hour) that we can detect with the specified test set up. Following Baarda and the more recent works by Imparato *et al.*, 2019 and Teunissen, 2006, the MDB can be determined by reversing the above procedure to calculate the detectability power γ and therefore fixing the type II error. Instead of calculating the integral of the probability density function of the alternative hypothesis, we fix $\gamma = 0.8$ (i.e. 80 % probability of correct detection), with now known model matrix $\mathbf{C}$ and covariance matrix $\mathbf{Q}_{yy}$. We check the value of the central normal distribution for the selected value of $\beta = 1-\gamma$, and use it to invert the equations, which leads to the MDB.

In a first step towards determining the MDB, we calculate the non-centrality parameter of the χ^2-distribution, $\lambda(\alpha, \beta, q)$ for significance level α and $\beta = 1-\gamma$, where γ is the detectability power of the test, and number of additional parameters q. In a first step the central χ^2-distribution is used to get the critical value k_α, that is the value where the integral over the central χ^2-distribution with degrees of freedom q reaches the value $1-\alpha$. Then we loop over all possible non-centrality values λ, to find $\lambda(\alpha, \beta, q)$, where the integral of the non-central χ^2-distribution for k_α as previously found and with degrees of freedom q reaches probability β.

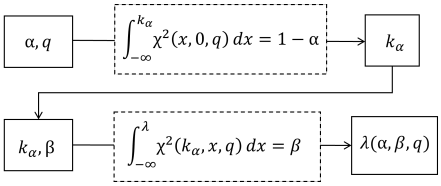

Figure 2.7: Schematic of the calculation of the non-centrality parameter $\lambda(\alpha, \beta, q)$. With $\chi^2(x, \lambda, q)$ we denote the probability density function of the non-central χ^2-distribution for (non-)centrality parameter λ and with degrees of freedom q.

The equation for the MDB for our case of known diagonal covariance matrix Q_{yy} reduces to to:

$$\text{MDB} = \sqrt{\frac{\lambda(\alpha, \beta, q)}{\mathbf{C}^T \mathbf{Q}_{yy}^{-1} \mathbf{C}}}. \tag{2.18}$$

For the respective alternative hypotheses defined above that simplifies to:

$$\text{MDB}_{step} = \sqrt{\frac{\lambda(\alpha, \beta, q)}{\sum_{i=\hat{t}}^{m} \sigma_i^{-2}}}, \tag{2.19}$$

indicating the minimum value of the step at location $\hat{t}$ for $H_{1,\hat{t}}$ and

$$\mathrm{MDB}_{lin} = \sqrt{\frac{\lambda(\alpha,\beta,q)}{\sum_{i=1}^{m} t^2/\sigma_i^2}}, \tag{2.20}$$

indicating the minimum slope value for a linear function for H_2.

As illustration, Figure 2.8 shows an example time series with MDB for two different models of alternative hypotheses: the height of a step function for a step after the 10th epoch and the slope of a linear trend starting at the first epoch. The step was assumed at a fixed location at $\hat{t} = 10$, without any further consideration of other alternative hypotheses (i.e. for different values of $\hat{t}$). We consider three different cases: Varying detectability power γ (Figure 2.8A), varying significance level α (Figure 2.8B) and in Figure 2.8C for both γ and α fixed and the same time series with larger uncertainty and lower uncertainty. It can be seen that larger detectability power γ and lower significance level α both lead to larger values of MDB. Additionally, higher standard deviation on the time series lead to larger values of MDB than lower standard deviations.

For comparison with the previously mentioned level of detection, this situation has been drawn for a simple 1-dimensional case in Figure 2.6. For the presented definition of the level of detection (see Section 2.2), the first test parameter to define is the significance level α, which determines the type I error and is mostly chosen at 5%. Here we show the level of detection with 95% confidence interval around the mean value y_0. The mean value for the alternative hypothesis y_1 is shown as the MDB when assuming $\gamma = 0.95$ (i.e. 95% probability of correct detection) in order to make them comparable. It can clearly be seen that y_1 is outside the level of detection region and the two terms do not coincide or provide the same insight.

2.5. RESULTS

2.5.1. EFFECTS OF ENVIRONMENTAL CONDITIONS ON ESTIMATION QUALITY

The effect of weather conditions on the height estimation quality was analysed in terms of mean elevation, before and after the tilt correction, as well as wind speed. Strong precipitation and fog, which both affect visibility, clearly have degrading effect on height estimations and most scans under these conditions will be excluded and not pass the quality criteria. In our entire data set, we find 156 cases, where point clouds were not available at the same time that fog or low visibility conditions were registered at the KNMI weather station at Schiphol airport. Schiphol is at about 25 km distance from the study site in Noordwijk, so the weather conditions could differ. But it gives an indication, that out of more than 5000 instances, where point clouds are not available, less than 3 % are possibly caused by low visibility.

The order of magnitude of the main influences on the uncertainty of the height estimate (z-coordinate) within the point cloud (on each grid cell) and on the entire point cloud are summarised in Table 2.3. We summarise, which effects are corrected or filtered out (scanner tilt and instrument failure) and how the others contribute to the two error terms σ_g and ε_{pc}. How temperature, wind speed and direct sunshine affects these uncertainty estimates is explained in more detail in the following subsections.

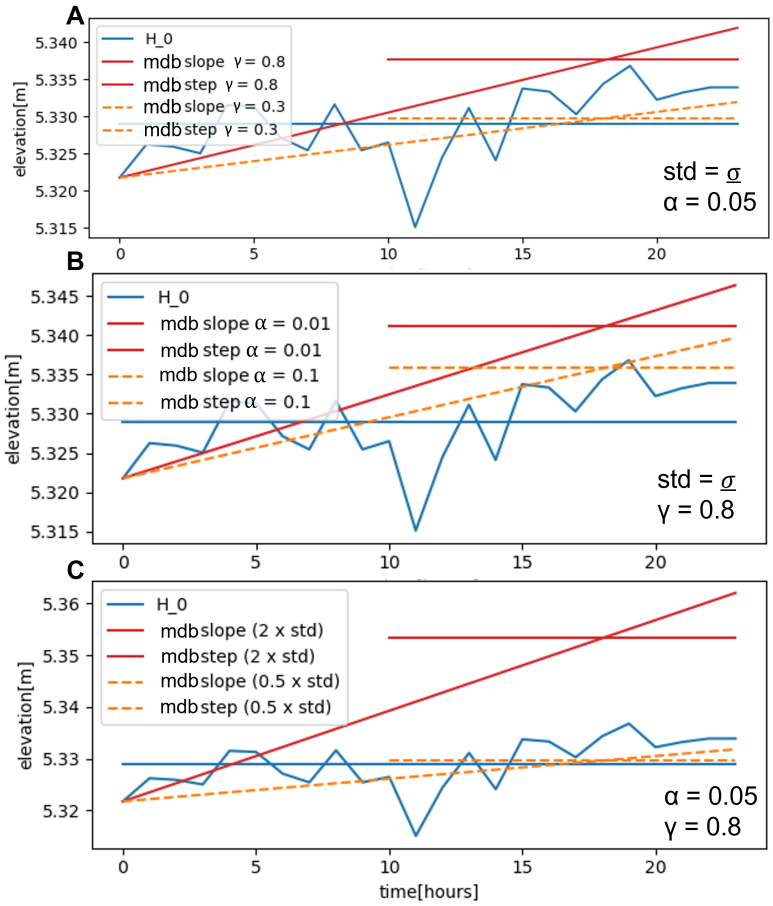

Figure 2.8: MDB for a time series covering 24 hours with two different alternative hypotheses: a step function, with a step at the 10th epoch and a linear function with positive slope. A: Both hypotheses are shown for different values of detectability power γ: 0.2 and 0.8. B: Both hypotheses are shown for different values of significance level α: 0.01 and 0.1. C: Both hypotheses are shown for fixed significance level and detectability power ($\alpha = 0.05$ and $\gamma = 0.8$) with different standard deviations: half standard deviation from the previous cases and doubled standard deviation.

TEMPERATURE EFFECTS

An example of the effect of temperature is shown for a period of six days in August 2019, with mostly clear skies, no precipitation and relatively low wind speed (average 4.5 m/s) in Figure 2.9. An obvious pattern in the mean elevation of the reference surface before correction is visible. The 24-hour pattern corresponds to duration of sunshine per

Influence	Order of magnitude	Explanation
Corrected and removed influences		
Scanner tilt	10 - 20 cm	Estimated from inclination angles and corrected
Instrument	≥ 10 cm	Errors due to instrument failure, affected point clouds are removed from analysis
Not corrected and considered in error terms ε_{pc} and σ_g		
Concrete expansion	0.7 cm	Estimated from temperature variations, not corrected, included in systematic error ε_{pc}
Registration	1.5 cm	Estimated from elevation of reference surfaces, not corrected, included in systematic error ε_{pc}
Atmosphere/temperature		
Surface roughness	0.9 - 7 cm	Estimated per grid cell, considered as random error σ_g
Surface moisture/ footprint size		Not quantified or corrected, included in random error σ_g

Table 2.3: Summary of estimated order of magnitudes for each of the influences on uncertainty affecting the entire point cloud or the grid cell in z-direction.

hour as well as temperature. Considering the entire six-day period, temperature as well as sunshine show negative correlation with the uncorrected mean elevation (-0.32 and -0.26 respectively). After correction, sunshine and mean elevation are not correlated anymore, but temperature and mean elevation now show a positive correlation of 0.28. When looking at the fast Fourier transformation (FFT) of the elevation, a slight 24-hour signal can be detected, which is probably the cause of the correlation with temperature. However, as can be seen in Figure 2.9, the main part of the 24-hour pattern in the elevation can be corrected for.

The remaining signal could be attributed to expansion of the concrete of the hotel building. An upward shift of the entire scanning set-up cannot be registered by the instrument's inclination sensors. The rotation correction is therefore not suitable to correct for this type of error. The concrete expansion, on days with highly varying temperatures, is estimated using the concrete expansion coefficient and the height of the building. It can amount to up to 0.7 cm elevation change, within several hours or days (not within the hour). However, we could not correlate the estimated concrete expansion directly with the remaining variation in elevation on the reference surfaces.

EFFECTS OF WIND SPEED

To show the effect of wind on the height estimation from PLS, we first visualise the combined pitch and roll angle (squared sum) together with wind speed, and show a clear

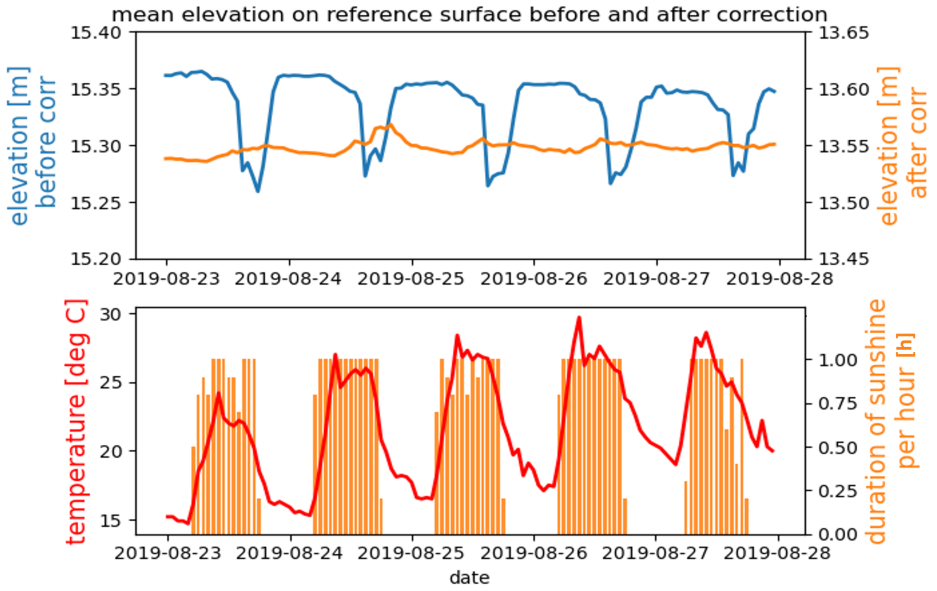

Figure 2.9: Time series of mean elevation over the course of six days in August 2019 on the helicopter landing platform before tilt correction (A) and after tilt correction (C) with duration of sunshine per hour (A) and temperature(B).

correlation between standard deviation of inclination values and wind speed (0.71) and a lower correlation between mean inclination and wind speed (0.47) for seven stormy days in February 2022, as an example. As was shown by Kuschnerus, Schröder, *et al.*, 2021, the effect of the higher standard deviation can be visible in the point clouds as a striped pattern on the reference surface. There, it was also shown that during strong winds, the sum of residuals on the reference surface is higher, and most likely related to the high standard deviation in the inclination values. For this example period in February 2022, we did not find a correlation between the sum of residuals on the reference platform and the inclination values or the wind speed.

Further, considering the entire data set, we did not find any correlation between wind speed and the residuals on the reference surface. This shows that the residuals in general are not that sensitive to wind and that the corrected point clouds are suitable for further analysis with the errors explained above. However, of the cases where there is a point cloud available, but the quality is not good enough, about 34 % occurred during strong winds (8 m/s and higher), which indicates a possible causal relation.

2.5.2. DIFFERENTIATING DYNAMIC PROCESSES ON THE BEACH

First, we show the MDB for test area 1 on the dry beach and two alternative hypotheses. Then, we perform the MHT (as introduced in Section 2.4.4) on both test areas, but for different length of time series and show the resulting partitioning of the areas according

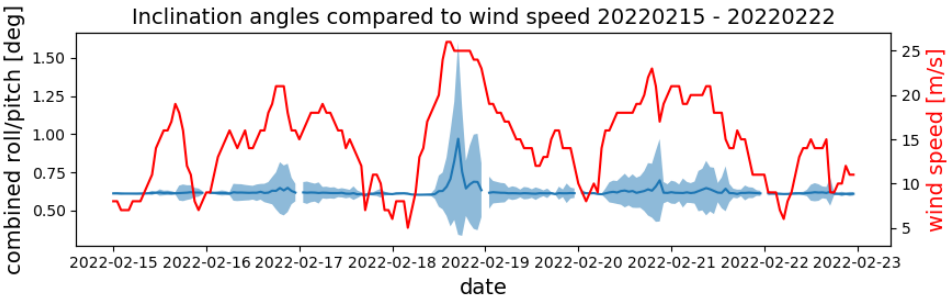

Figure 2.10: Combined inclination angles (squared sum of pitch and roll) with standard deviation and wind speed.

to the most likely estimated model.

MINIMAL DETECTABLE BIAS

We show the MDB for test area 1 as indicated in Figure 2.3: On the dry beach, next to the beach cafe and frequently used by people as well as effected by bulldozer works. We estimate the MDB as explained in Section 2.4.6 with detectability power $\gamma = 80\%$ and with significance level $\alpha = 95\%$.

As an example to illustrate the MDB on the dry test area we use 24-hour time series from 7 of January 2020. The MDB is calculated for a sudden change in form of a step function happening at 17:00h in the afternoon and shown in Figure 2.11. We know that bulldozer works started that day and moved a considerable amount of sand between 11:00h and 17:00h depending on location, as reported previously (Kuschnerus, Lindenbergh, Lodder, *et al.*, 2022). The MDB shows, what the minimum change in elevation per grid cell would have to be at that time, in order to be detected as significant. We incorporate the systematic error ε_{pc} in order to derive a realistic estimate. The resulting MDB ranges from 1.8 cm (just above the ε_{pc}) up to 32 cm for some outliers, with a median at 3.1 cm. As comparison we show the LoD as defined above, which has a median value of 26.8 cm for the entire area, which is one order of magnitude higher than the MDB with our configuration. The MDB and the average value of the LoD indicated in Figure 2.16, are representative of one dot in Figure 2.11A and C respectively. The MDB for a linear trend was calculated as well: Only a few grid cells appear to deviate a lot from the median slope of 3.2 cm/day, see Figure 2.11B.

HYPOTHESIS TESTING

We applied the MHT as explained in Section 2.4.4 on the time series of each grid cell in the previously described test areas, using 24-hour time series on 7 January 2020.

Using the above described testing method for the test area on the dry part of the beach we identify sudden changes, most likely caused by bulldozer works, as mentioned before. The changes happen between 16:00 and 17:00 in the afternoon and are identified

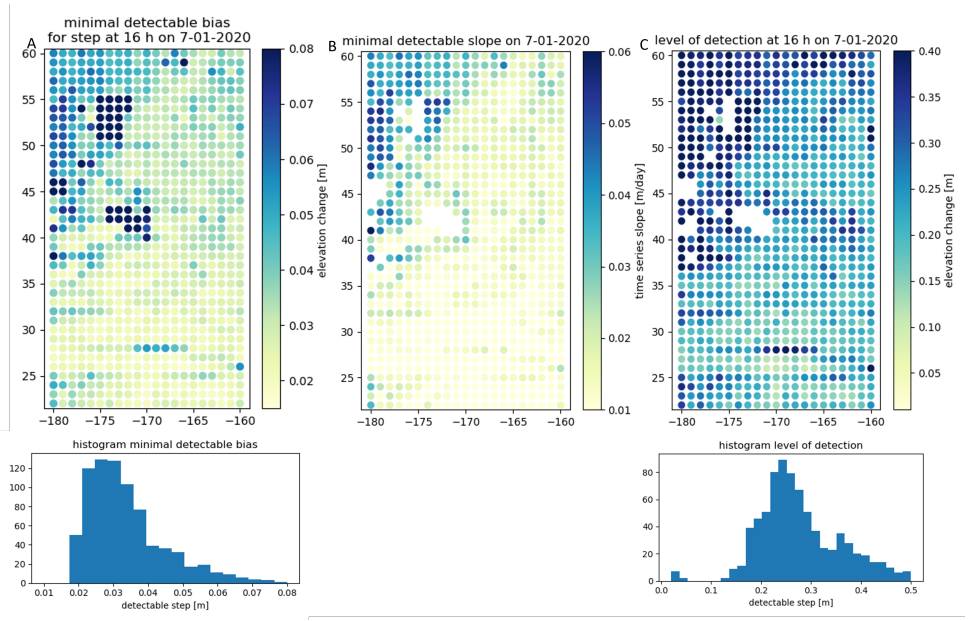

Figure 2.11: MDB at significance level $\alpha = 0.05$ and with detectability power $\gamma = 0.8$ for a step function with step at 17:00h (A) and a linear trend (in m/day) (B). Results were calculated for each 24-hour time series on 7 January 2020. For comparison the level of detection is shown for a significant change at 17:00h (C). To simplify the comparison of the spread of the values for the MDB and LOD the histograms of the respective estimates are shown as well.

on a large part of the area (see Figure 2.12). Figure 2.12 shows some example time series with no significant change and sudden changes.

Similar to results shown in previous works (Barbero-García *et al.*, 2023; Kuschnerus, Lindenbergh, Lodder, *et al.*, 2022; Kuschnerus, Lindenbergh, & Vos, 2021), human activities on the beach like bulldozer works that lead to sudden elevation changes can be detected and quantified. MHT for 24-hour elevation time series with a step function as alternative hypothesis allows to find and identify most bulldozer works taking place on the sandy beach.

To identify longer term and slower processes, we apply the same testing procedure to time series in the same area, covering the entire month of July 2021, with two epochs per day, at noon and at midnight. Again we can classify the testing area into the three classes: no change, sudden change and linear trend. Here we find more areas with a linear change. However, for some cases a two-(or more)step model would seem more appropriate, as can be seen in Figure 2.13B. The sudden change model does not fit the time series very well, since in a longer term time series, it is more likely that more than one sudden change occurs (see Figure 2.13 C).

In Figure 2.13B slow erosion processes are identified. The area close to the dune foot

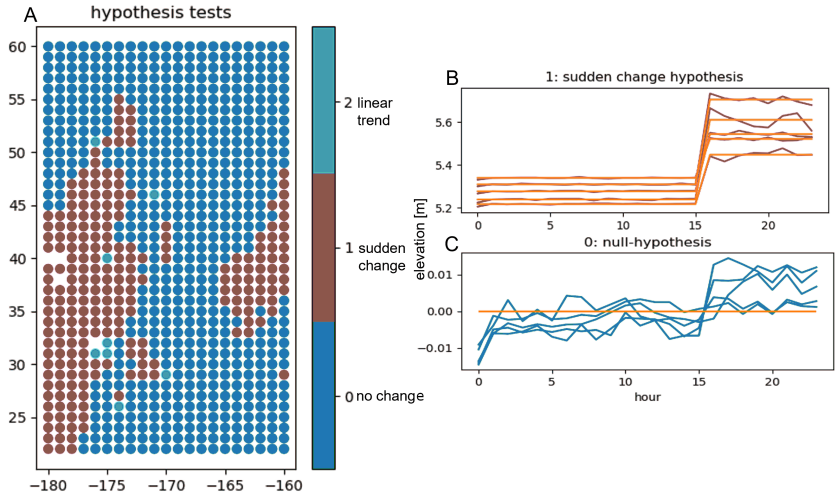

Figure 2.12: A: Dry area classified according to hypothesis testing based on time series covering 24 hours on 07-01-2020. The x- and y-axis represent across-shore and along-shore distance in meters. B: Example time series with sudden change hypothesis sustained. The changes happened at 16:00h in the afternoon. C: Example time series over 24 hours with null-hypothesis not rejected, i.e. no statistically significant change detected. The small step occurring in some time series around 16:00h is not significant enough to be detected.

shown in panel A of the same figure is eroding with about 0.5 cm per day. A slow, gradual erosion process like this can be caused by aeolian sand transport or possibly by frequent use of beach visitors crossing that area. It is however unique to observe the gradual changes over a long period of time and quantify the erosion at sub-centimetre level.

When considering test area 2 (as indicated in Figure 2.3) in the intertidal area of the beach, fully populated daily time series are not available. The area appears in the point clouds only once (or sometimes twice) per day, during low tide. A set of daily time series in September 2020, with one epoch per day, chosen at low tide, where the area of non-empty grid cells in the DEM is largest, was analysed using MHT. Figure 2.14 shows some examples of the classified time series (B) and an overview of the entire area coloured with the most likely assumed hypothesis (A). Here we distinguish between erosion (negative) and deposition (positive) versions of the found model time series, to give more insight in the present dynamics.

It becomes apparent that the two models used for the alternative hypotheses are not sufficient to classify all processes affecting the beach. The model library can be extended with other models and additionally the start and stop time of the time series that is classified should be chosen individually per time series in order to fit the available models better. Still a general trend of either eroding or accreting becomes apparent in the

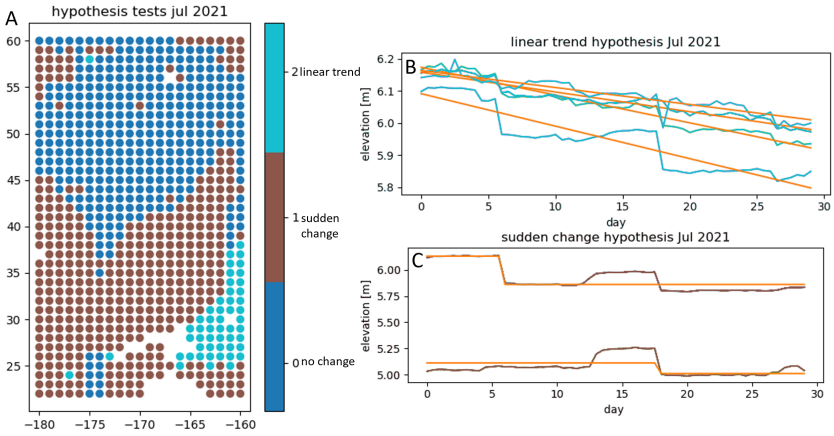

Figure 2.13: A: Dry area classified according to hypothesis testing based on time series covering the month of July 2021. The x- and y-axis represent across-shore and along-shore distance in meters. B: Example time series with linear trend (orange). One of the time series (in blue) shows a shape that appears more like two sudden changes. C: Two examples, where the sudden change was found as most likely model (orange), but does not fit the shape of the time series, which shows more than one sudden change.

overview plot (Figure 2.14A) and the more dynamic nature of the intertidal area is clearly visible. The observed effects are most likely the result of tides and waves depositing and eroding sand while the area is flooded every day.

2.6. DISCUSSION

2.6.1. ADDITIONAL EFFECTS ON ESTIMATION QUALITY

Since the laser scanner is pointed to the west, the sun could directly shine into the lens of the scanner in the hours just before sunset. Since we do not have weather data from directly next to the laser scanner for most of the time, there is no way to verify when this situation occurred. We collect all point clouds, acquired a few hours before and up to the time of sunset on that day and check if they pass our quality checks. All point clouds with low quality collected just before or at sunset could potentially be affected by direct sun shine. We filtered out all instances with high wind speeds (above 8 m/s) and precipitation (more than 1 mm/h). There are 10 instances, where around the time of sunset, the quality of the DEM was not good enough (i.e. the residuals on the reference surface were too high, compare threshold defined in section 2.4.1) during or in the hours before sunset. All of them occurred in the summer months of 2019, outside of the period of malfunction of the laser scanner. The respective point clouds show some noise (i.e., random points in the air), but the cause of this noise could not be determined with certainty. We observed a similar effect in the weeks leading up to the maintenance of the scanner

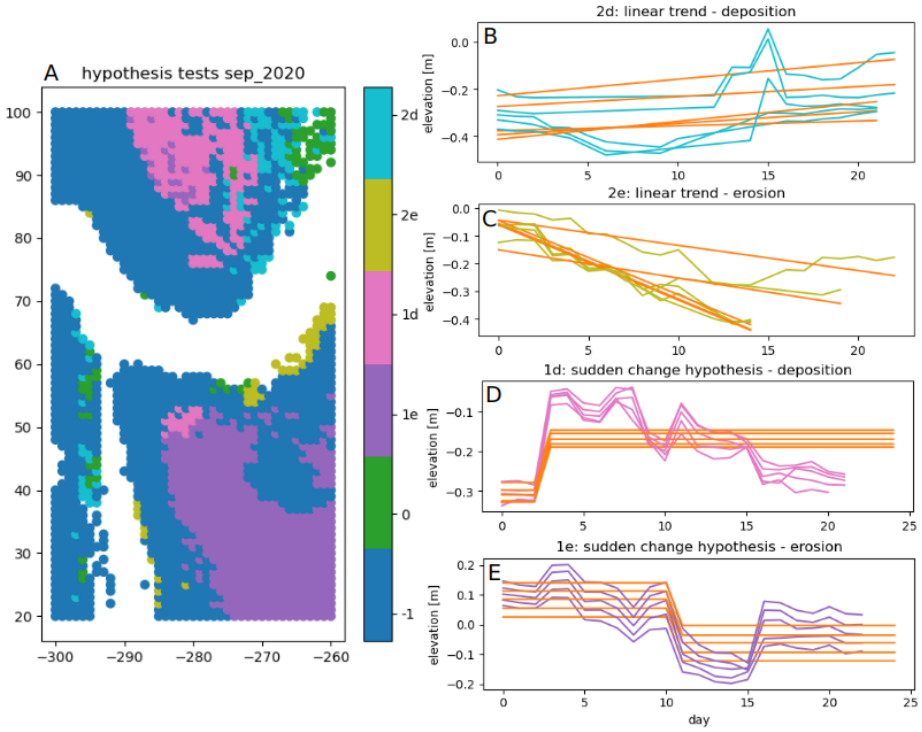

Figure 2.14: A: Classification according to most likely assumed hypothesis in the intertidal area, based on daily time series in September 2020. The x- and y-axis represent across-shore and along-shore distance in meters. For grid cells in green (label 0) the null-hypothesis was the most likely. Blue grid cells (label -1) indicate locations where no model could be found to match the time series. B-E: Example time series for each of the four assumed models: positive linear trend (deposition, B), negative linear trend (erosion, C), sudden change deposition (D) and sudden change erosion (E). This example illustrates, that the two models used for the alternative hypothesis are not sufficient for many cases and an adaptive length of the time series would improve the results.

in spring 2020, when the general quality of height estimations declined due to the previously mentioned instrument malfunction. With a ray-tracing method as for example presented by Zhou *et al.*, 2019 the times with direct sunshine into the laser scanner could be determined with more certainty, which would allow more investigation into possible effects on the respective point clouds.

Beach visitors show up as peaks or outliers in the grid cell time series. As shown in Figure 2.15, the hourly time series during a scan by students of the TU Delft show peaks whenever a person was present in the respective grid cell. These outliers are not causing

the null-hypothesis to be rejected, but a busy period with lots of beach visitors will lead to higher grid cell error σ_g, which has an impact on the hypothesis testing.

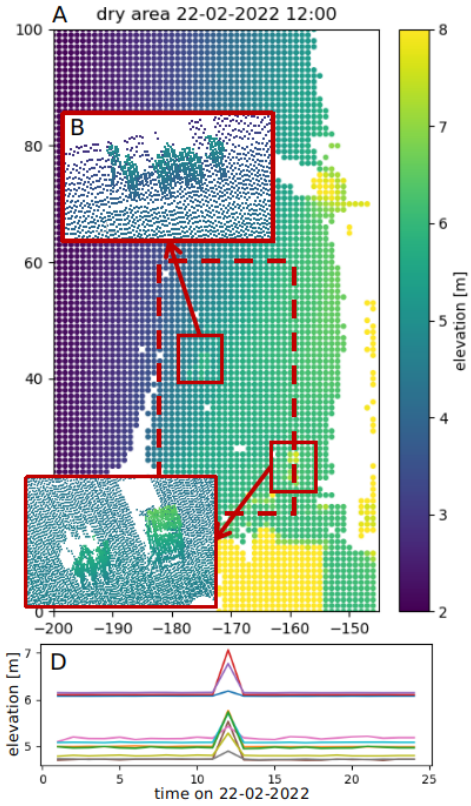

Figure 2.15: Effect of beach visitors on the DEM. On 22-02-2022, a group of students visited the study site at the beach and appears in the point cloud at 12:00. The DEM has grid cells of about 80 cm higher than the surroundings at that time (A). The corresponding point clouds showing the people (B,C). The corresponding time series of the entire day show a clear outlier at 12:00 (D).

A possible solution to this would be a filter for dynamic objects on the beach applied to each point cloud before generating the DEMs. This would increase processing times significantly and potentially lead to some data loss, but in return it could make the results of MHT more reliable.

2.6.2. POTENTIAL IMPROVEMENTS OF SET-UP AND PROCESSING

For this research we have made several assumptions and simplifications considering the atmospheric conditions and calibration of the instrument. Additionally we did not consider in detail the slope of the beach, footprint size variation, surface roughness, fine

registration and correlation of the height estimates.

For a future improved set-up we recommend installing temperature sensors or even entire weather stations next to the scanner and in the observed scene at various locations. The collected temperature measurements would allow for an estimation of the refraction index and other atmospheric influences. To improve atmospheric correction with the additional atmospheric measurements we suggest to consider the approaches of Voordendag *et al.*, 2023 and Czerwonka-Schröder, 2023.

Any instrumental error could potentially be further reduced by performing a dedicated calibration in regular time intervals on site, preferably without removing the instrument from its permanent location. This could reduce any drifts in range estimates and yaw alignment and prevent instrument failure and need for maintenance, or at least allow earlier detection leading to reduced data loss. For calibration of terrestrial laser scanners one could consider for example the works of Medic *et al.*, 2019 and Schmitz *et al.*, 2019.

The slope of the beach and with that the incidence angle and their variations through time could be estimated based on point clouds or several neighbouring DEM grid cells. Then, a slope correction could be applied to the individual height estimates. The same holds for the estimation of footprint sizes per grid cell. A mean value for footprint size depending on location and estimated incidence angle could be taken into account for the individual height estimates.

The contribution of the surface roughness to the error per grid cell of up to 7 cm is quite high. There are grid cells on the beach, where a constant elevation is not an appropriate surface model. Examples include the edge of a cliff, deep tire tracks, or channels in the intertidal zone. To take such topography into account one could apply different methods of surface approximation, as for example presented by Kermarrec *et al.*, 2022.

The entire data set could be subject to a fine registration method, aligning all point clouds to a suitable epoch, using for example the iterative closest point (ICP) method. This is a computationally intensive operation, but has been done on a similar data set for example by Vos *et al.*, 2022.

Finally, we did not consider correlation between consecutive height estimates. When dealing with short-term measurements, as in our case the hourly data set, temporal correlation between the individual height estimates is likely. To incorporate a correlation coefficient into the hypothesis testing model, the theory of Baarda can be applied, as for example discussed in Rofatto *et al.*, 2020.

2.6.3. COMPARISON TO EXISTING METHODS

Other approaches developed by Anders *et al.*, 2020; Anders *et al.*, 2021 and Kuschnerus, Lindenbergh, and Vos, 2021 make use of region growing and clustering to extract regions in space and time with similarly behaving time series of elevation (or elevation differences). These methods have the advantage that it does not have to be known previously, what kind of processes are expected, so they can in principle be applied without pre-existing knowledge on the occurring changes. Winiwarter *et al.*, 2023 combine Kalman filtering for time series interpolation, level of detection and clustering on extracted features to find dynamic processes that were not previously defined. Also here the authors avoid using prior process knowledge. The level of detection is used here for the comparison of height estimates at each epoch with a reference height. For illustration of the

typical use, we consider a 24-hour time series on January 7th, 2020 and show all differences to the first epoch, together with the level of detection (blue), see Figure 2.16. We use equation 2.1 with $\varepsilon_{pc} = 0.015$ m as registration error for the calculation of the level of detection. We can therefore derive, at what point the time series deviates from the first epoch with 95% confidence (first type error). We further indicate the MDB for both curves and significance level $\alpha = 95\%$ and detectability power $\gamma = 80\%$.

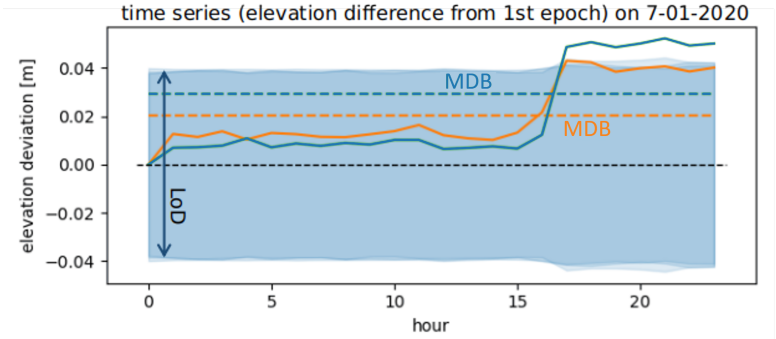

Figure 2.16: Example of level of detection for two 24-hour time series on 7 January 2020. The time series are shown as difference in elevation from the first epoch. The level of detection is indicated in blue and the MDB for a step function for both time series is shown as a dashed line. The MDB and LoD shown here are different representations of one dot in Figure 2.11 A and C, respectively.

None of these methods are specifically taking into account the multi-epoch nature of our data, nor do they efficiently detect processes, affected areas and the nature and/or cause of the process. The here presented method uses previous knowledge on expected processes, in order to determine the temporal sampling, as well as the models for the alternative hypothesis. In this way specific processes can be detected together with a provided level of confidence. To make this usable for the entire observed area, and more for example periodic processes, the model catalogue can be extended and possibly a combination of the above mentioned clustering methods with MHT could be used to identify relevant areas and/or time intervals in the three-year observation period.

We can see in Figures 2.12 and 2.14 that the MHT procedure as applied in this research allows a classification of time series into different processes and therefore also a grouping according to process. The latter has the potential to be used in a similar way or complementary to other methods of clustering time series like Kuschnerus, Lindenbergh, and Vos, 2021 or Winiwarter *et al.*, 2023. The main difference with these methods is the pre-definition of the expected alternative hypotheses. By defining the alternative hypotheses we determine apriori, which processes we are expecting to find. Our search is limited to those processes, but on the other hand provides a level of confidence when detecting them.

For the hypothesis tests the model library can potentially be extended in order to fit more processes present on the beach. More simple functions can easily be added, as

for example a periodic elevation change, step functions with several steps or a combination of linear trends and step functions. However, the difficulty in the application will lie in matching the models and the relevant time scales. Considering hourly time series over one day, our simple three model hypothesis testing showed good results in identifying anthropogenic changes, however it did not categorise the intertidal area on monthly time scales well. On larger time scales, other additional models would be required. Therefore, again a priori knowledge on the length and time scale of the time series and the relevant, expected processes is needed in order to identify geomorphological coastal processes in a 4D point cloud data set.

2.6.4. ASSUMPTIONS ON UNCERTAINTY

Implicitly we make the assumption that larger residuals per grid cell lead to more uncertainty in the knowledge of the mean elevation at that epoch. This assumption holds in many cases, but is difficult to verify in general. There could be cases, for example in the intertidal area, where surfaces are relatively smooth, but with sharp edges, where this assumption is not true. Another case could be a frozen beach during temperature below zero, where the surface roughness stays frozen over several epochs, thus it does not effect the estimation of the model as we are assuming. However, these cases are not generally true, and we assume the chosen approach provides a realistic estimation of the variance in most cases.

We also tested to use a rolling-window standard deviation based purely on the elevation time series, discarding the spatial component/surface roughness but incorporating the temporal domain into the standard deviation. This did not lead to more realistic assumptions on the standard deviation. The temporal component leads to a lower standard deviation (compared to surface roughness) and is highly sensitive to jumps/sudden changes in the time series. The assumption that a sudden change in elevation made our measurements less likely does not hold in most cases and we therefore conclude that the spatial (roughness-based) standard deviation is a better approximation of the true measurement uncertainty.

2.7. CONCLUSION AND RECOMMENDATIONS

We investigated the main effects on height estimates of a 4D point cloud data set from PLS and applied MHT as well as the MDB to time series of height estimates from these observations answering the following questions.

What is the minimal change in height on a sandy beach that can be estimated with a given confidence with our permanent laser scanning set-up? The MDB is suitable for the estimation of possible elevation changes found within time series from PLS. Making use of the long elevation time series available in the PLS data set, the parameters for the estimation of the MDB can easily be tuned to detect small scale changes in sandy beach elevation with a defined detectability power and to identify different models of elevation change. This provides advantages over the commonly used level of detection, especially the known detectability power and the considering of entire elevation time series rather than just pairs of epochs. In our example case, slopes of 0.032 m/day and

sudden changes of 0.031 m were identified with statistical power of 80% and with significance 95% in 24-hour time series on the upper beach.

How do environmental conditions contribute to the uncertainty of height estimates from permanent laser scanning for the identification of change processes? The main effects on height estimates from PLS are weather conditions such as strong wind causing the instrument to shake and precipitation and instrument malfunctions. When these are filtered out the remaining effects are dominated by the surface roughness and registration error, which we estimate as below 7 cm and 1.5 cm, respectively. Height deviations caused by temperature changes can reach up to 1 cm within a day and are largely corrected by the application of a rotation matrix based on inclination estimates from the scanner's internal inclination sensor.

Which change processes on a sandy beach can be observed with permanent laser scanning and at which temporal and spatial scales? MHT allows for the grouping of areas following similar processes. At the same time we are classifying each time series according to the available alternative hypotheses and therefore gain more insight into the predominant change regime. It is especially suitable to detect and quantify slow longer term erosion and accretion processes, which are most likely caused by aeolian sand transport and difficult to observe using other techniques

The presented methods appear to be promising in the large scale processing and data mining for change processes within a 4D data set from PLS. To be applicable to entire elevation time series of length longer than one month, a method to segment each time series and detect break points would allow to treat each segment separately. We could then apply MHT per segment with appropriate models for the alternative hypothesis.

Further investigation of the error sources on the height estimates and quantification of each effect could potentially improve the quality and reliability of height estimates and therefore lead to the detection of even smaller elevation changes and trends in elevation.

The presented method could potentially be adapted for different applications, such as height estimation on glaciers from PLS or identification of height changes in vegetation estimated from laser scanning observations.

3

Coastal Change Patterns from Time Series Clustering of Permanent Laser Scan Data

Mieke Kuschnerus, Roderik Lindenbergh, Sander Vos

Sandy coasts are constantly changing environments governed by complex, interacting processes. Permanent laser scanning is a promising technique to monitor such coastal areas and to support analysis of geomorphological deformation processes. This novel technique delivers 3D representations of the coast at hourly temporal and centimetre spatial resolution and allows to observe small scale changes in elevation over extended periods of time. These observations have the potential to improve understanding and modelling of coastal deformation processes. However, to be of use to coastal researchers and coastal management, an efficient way to find and extract deformation processes from the large spatio-temporal data set is needed. To enable automated data mining, we extract time series of surface elevation and use unsupervised learning algorithms to derive a partitioning of the observed area according to change patterns. We compare three well known clustering algorithms, k-means, agglomerative clustering and DBSCAN, apply them on the set of time series and identify areas that undergo similar evolution during one month. We test if these algorithms fulfil our criteria for suitable clustering on our exemplary data set. The three clustering methods are applied to time series over 30 days extracted from a data set of daily scans covering about two kilometres of coast at Kijkduin, the Netherlands. A small section of the beach, where a pile of sand was accumulated by a bulldozer is used to evaluate the performance of the algorithms against a ground truth. The k-means algorithm and agglomerative clustering deliver similar clusters, and both allow to identify a fixed number

This chapter has been published in Earth Surface Dynamics **9**, 89–103 (2021) Kuschnerus, Lindenbergh, and Vos, 2021.

of dominant deformation processes in sandy coastal areas, such as sand accumulation by a bulldozer or erosion in the intertidal area. The level of detail found with these algorithms depends on the choice of the number of clusters k. The DBSCAN algorithm finds clusters for only about 44% of the area and turns out to be more suitable for the detection of outliers, caused for example by temporary objects on the beach. Our study provides a methodology to efficiently mine a spatio-temporal data set for predominant deformation patterns with the associated regions, where they occur.

3.1. INTRODUCTION

Coasts are constantly changing environments that are essential to the protection of the hinterland from the effects of climate change and, at the same time, belong to the areas that are most affected by it. Especially long-term and small scale processes prove difficult to monitor but can have large impacts (Aarninkhof *et al.*, 2019). To improve coastal monitoring and knowledge of coastal deformation processes, a new technique called Permanent Laser Scanning (PLS) (also called continuous laser scanning) based on LiDAR measurements is available. For this purpose, a laser scanner is mounted on a high building close to the coast in a fixed location acquiring a 3D scan every hour during several months up to years.

The resulting spatio-temporal data set consists of a series of point cloud representations of a section of the coast. The high temporal resolution and long duration of data acquisition in combination with high spatial resolution (in the order of centimetres) provides a unique opportunity to capture a near continuous representation of ongoing deformation processes, like for example storm and subsequent recovery, on a section of the coast. As reported by Lazarus and Goldstein, 2019, the natural effects of a storm on a typical urban beach can rarely be analysed separately from anthropogenic activities, since in most cases work with bulldozers starts immediately after or even during severe storms. There is a need for the detection and quantification of change processes that influence the geomorphology of the coast, to allow understanding and modelling them, as the reaction of the coast to extreme weather events gains importance (Masselink & Lazarus, 2019). More examples for potential use of such a data set are presented by O'Dea *et al.*, 2019, who use data from a similar set-up in Duck, USA.

The PLS data set is large (in the order of hundreds of gigabytes), and to be relevant, the information on deformation processes has to be extracted concisely and efficiently. Currently there are no automated methods for this purpose and studies focus on one or a few two dimensional cross-sections through the data (for example O'Dea *et al.*, 2019). The high temporal resolution and long observation period lead to a high dimensional data set of long time series (i.e. 30 data points up to several thousands). Data mining on high dimensional data sets can be challenging as discussed for example by Zimek *et al.*, 2012. In a first step towards extraction of interesting events and change patterns we build on the method introduced by Lindenbergh *et al.*, 2019. We use clustering algorithms on time series representing the evolution of topography, to group these time series according to their similarity in change pattern and then identify underlying processes. We use clustering (or unsupervised learning) to avoid having to specify the patterns and processes that we are looking for in advance.

One example of spatio-temporal segmentation on our data set from PLS was recently developed by Anders *et al.*, 2020. They detected seed points for deformation in time series from PLS, to then grow a region affected by the detected change around the seed points with the use of dynamic time warping distance to spatial neighbours. Dynamic time warping is a distance measure between time series, that accounts for similarity in patterns even though they might be shifted in time (see for example Keogh and Ratanamahatana, 2005). One drawback of this approach is that temporal patterns of interest have to be defined before hand, and therefore only deformation patterns that are expected can be found. Another approach to model spatio-temporal deformations in point clouds

from laser scanning, is presented by Harmening and Neuner, 2020. Their model assumes that the deformation can be represented by a continuous B-spline surface. This approach could potentially be used to further analyse some of the deformation patterns found in our study but does not allow the exploratory data mining, that we are aiming to accomplish. A more general overview of methods to find spatio-temporal patterns in earth science data was published by Tan *et al.*, 2001 and a continuation of this study was presented by Steinbach *et al.*, 2001. The study of Tan et al. deals with preprocessing of time series of different variables from satellite data including issues with auto-correlation and seasonality. Steinbach et al. successfully apply a novel clustering technique introduced by Ertöz *et al.*, 2003 to explore spatio-temporal climate data. However, this technique only focuses on contiguous clusters, where all time series are in a close neighbourhood to each other, and does not allow to find general patterns independent of location.

Time series data sets are also used to asses patterns of agricultural land use by Recuero *et al.*, 2019. For this study time series of Normalized Difference Vegetation Index (NDVI) data have been analysed using auto-correlation values and random forest classification. Benchmark data from an alternative source was needed to train the classifier. Such benchmark data is currently not available in our case. A study by Belgiu and Csillik, 2018 used time series from Sentinel-2 satellite data for cropland mapping. They made use of dynamic time warping classification and showed that in areas with little available reference data for training a classifier, their approach delivers good results in segmentation based on time series' evolution. Also in this case pre-labelled training data is required. Another approach using expectation-based scan statistics was presented by Neill, 2009: To detect spatial patterns in time series from public health data, a statistical method based on expectation values is used. Clusters are formed where the observed values significantly exceed the expectation. The results are promising but depend on the choice of time series analysis method, statistics used and the shape of the search region, which all have to be defined in advance specific to each data set and application. Generally there is a lack of studies on mining spatio-temporal data for deformation patterns, without using training data or predefined change patterns.

The goal of the present study is to evaluate the application of clustering algorithms on a high dimensional spatio-temporal data set without specifying deformation patterns in advance. Our objectives in particular are:

1. To analyse and compare the limits and advantages of three clustering algorithms for separating and identifying change patterns in high dimensional spatio-temporal data.

2. To detect specific deformation on sandy beaches by clustering time series from permanent laser scanning.

We compare the k-means algorithm, agglomerative clustering and the DBSCAN algorithm on a PLS data set over 30 days, to investigate the effectiveness of the identification of coastal change patterns. All three algorithms are well established and represent three common but different approaches to data clustering. To determine if an algorithm is suitable, we expect that it fulfils the following criteria:

- A majority of the observation area is separated into distinct regions,
- each cluster shows a change pattern that can be associated with a geomorphic deformation process, and
- time series contained in each cluster roughly follow the mean change pattern.

We use the different clustering approaches on a small area of the beach at the bottom of a footpath, where sand accumulated after a storm, and a bulldozer subsequently cleared the path and formed a pile of sand. We determine the quality of the detection of this process for all three algorithms and compare them in terms of standard deviation within the clusters and area of the beach covered by the clustering. We compare and evaluate the resulting clusters using these criteria as a first step towards the development of a method to mine the entire data set from PLS for deformation processes.

3.2. THE PERMANENT LASER SCAN DATA SET

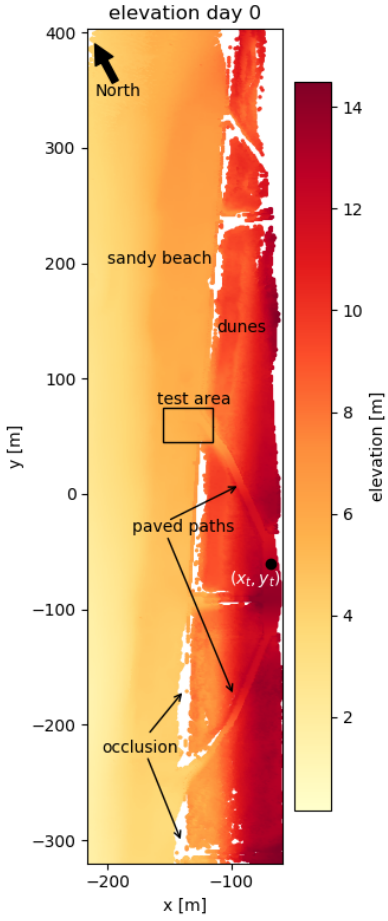

Figure 3.1: Top view of a point cloud representing the observation area at low tide on 1st January 2017. The laser scanner is located at the origin of the coordinate system (not displayed). The point (x_t, y_t) indicates the location of the time series shown as an example in Figure 3.3. The test area, which is discussed in Section 3.3.4, is indicated with a box at the end of the northern path leading to the beach. The paved paths leading to the beach are used as stable reference surface for the errors reported in Table 3.1. Parts that are white between the dunes and the sandy beach are gaps in the data due to occlusions caused by the dunes.

The data set from PLS is acquired within the CoastScan project at a typical urban beach in Kijkduin, the Netherlands, Vos *et al.*, 2017. For the acquisition a Riegl VZ-2000 laser scanner was used to scan over a period of six months from December 2016 to May 2017. The full data set consists of hourly scans of a section of sandy beach and dunes.

Figure 3.2: Riegl VZ2000 laser scanner mounted on the roof of a hotel facing the coast of Kijkduin, the Netherlands. The scanner is covered with a protective case to shield it from wind and rain.

For the present study, a subset of the available data is used to develop the methodology. This subset consists of 30 daily scans taken at low tide over a period of one month, January 2017. It covers a section of the beach and dunes in Kijkduin and is displayed in top view in Figure 3.1. The area contains a path and stairs leading down to the beach, a paved area in front of the dunes, a fenced in dune area and the sandy beach. It is about 950 m long, 250 m wide and the distance from the scanner to the farthest points on the beach is just below 500 m. For the duration of the experiment the scanner was mounted on the roof of a hotel just behind the dunes at a height of about 37 m above sea level (as shown in Figure 3.2).

The data is extracted from the laser scanner output format and converted into a file that contains xyz-coordinates and spherical coordinates for each point. The data is mapped into a local coordinate system, where the origin in x- and y-direction is at the location of the scanner and the height (z-coordinate) corresponds to height above sea level. Since we are interested in relative changes between consecutive scans, we do not transform the data into a geo-referenced coordinate system for this analysis.

Each point cloud is chosen to be at the time of lowest tide between 18:00 and 06:00, in order to avoid people and dogs on the beach, with the exception of two days where only very few scans were available due to maintenance activities. The data from 9th of January 2017 is entirely removed from the data set, because of poor visibility due to fog. This leads to the 30 day data set, numbered from 0 to 29. Additionally all points above 14.5 m elevation are removed to filter out points representing the balcony of the hotel

and flag posts along the paths. In this way also a majority of reflections from particles in the air, birds or raindrops are removed. However, some of these particles might still be present at lower heights close to the beach.

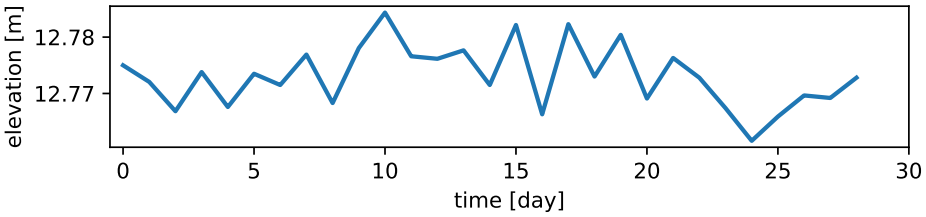

Figure 3.3: Time series of elevation at location (x_t, y_t) (marked in Figure 3.1) on the path that is assumed to be stable throughout the entire month. Elevation is varying within less than 2 cm.

Since the data is acquired from a fixed and stable position we assume that consecutive scans are aligned. Nevertheless, the orientation of the scanner may change slightly due to strong wind, sudden changes in temperature, or maintenance activities. The internal inclination sensor of the scanner measures these shifts while it is scanning and we apply a correction for large deviations (more than 0.01 degrees) from the median orientation.

The remaining error in elevation is estimated as the standard error and the 95-percentile of deviations from the mean elevation over all grid cells included in the stable paved area. We chose the stable surface that is part of the paved paths on top of the dunes and leading to the beach in northern and southern direction as indicated in Figure 3.1. This area includes 1653 grid cells with complete time series. The derived mean elevation, standard error and overall 95-percentile of deviations from the mean per time series averaged over the stable area are reported in Table 3.1. The elevation does on average not deviate more than 1.4 cm from the mean elevation, and 95 % of deviations from the mean elevation are on average below 3.5 cm. An example time series from the stable paved area on top of the dunes (at location (x_t, y_t) marked in Figure 3.1) is shown in Figure 3.3.

Table 3.1: Test statistics of the gridded elevation values on the paved area, which is assumed to be stable throughout the observation period of one month. Values are calculated per time series and averaged over the entire stable area, which results in mean elevation, standard error and an average 95-percentile of deviations from the mean.

mean elevation	12.43 m
standard error	1.4 cm
95-percentile of deviation from mean (averaged over all grid cells)	3.5 cm

3.3. METHODS

To derive coastal deformation processes from clusters based on change patterns we follow three steps: Extraction of time series, clustering of time series with three different algorithms, and derivation of geomorphological deformation processes. To cluster time series the definition of a distance between two time series (or the similarity) is not immediately obvious. We discuss two different options (Euclidean distance and correlation) to define distances between time series with different effects on the clustering results. The rest of this section is organized as follows: We focus on time series extraction in subsection 3.3.1, discuss distance metrics for time series (3.3.2), introduce three clustering algorithms (3.3.3) and our evaluation criteria (3.3.4). The derivation of deformation processes will be discusses with the results (section 3.4).

3.3.1. TIME SERIES EXTRACTION

Time series of surface elevation are extracted from the PLS data set by using a grid in Cartesian xy-coordinates. We only use grid cells that contain at least one point for each of the scans.

Before defining a grid on our observed area, we rotate the observation area to make sure that the coastline is parallel to the y-axis, as shown in Figure 3.1. This ensures that the grid covers the entire observation area efficiently and leaves as few empty cells as possible. Then we generate a regular grid with grid cells of 1 m × 1 m. Time series are generated for each grid cell by taking the median elevation z_i for each grid cell and for each time stamp t_k. That means, per grid cell with centre (x_i, y_i) we have a time series

$$\tilde{Z}_i = (z_i(t_1), \dots z_i(t_T)), \tag{3.1}$$

with the number of time stamps $T = 30$. To make the time series dependent on change patterns, rather than the absolute elevation values, we remove the mean elevation $\bar{z}_i$ of each time series $\tilde{Z}_i$. This leads to time series

$$Z_i = (z'_i(t_1), \dots, z'_i(t_T), \tag{3.2}$$

with $z'_i(t_k) := z_i(t_k) - \bar{z}_i$.

In this way we extract around 40 000 grid cells that contain complete elevation time series for the entire month. The point density per grid cell varies depending on distance to the laser scanner. For example, a grid cell on the paved path (at about 80 m range) contains about 40 points (i.e. time series at (x_t, y_t) in Figure 3.1), whereas a grid cell located close to the water line, at about 300 m distance from the scanner, may contain around three values. This implies that the median per grid cell is based on more points the closer a grid cell is to the scanner.

3.3.2. DISTANCE METRICS

We consider two different distance metrics for our analysis: the Euclidean distance as the default for the k-means algorithm and agglomerative clustering, and correlation distance for the DBSCAN algorithm.

EUCLIDEAN DISTANCE

The most common and obvious choice is the Euclidean distance metric defined as:

$$d_E(Z_0, Z_1) = ||Z_0 - Z_1|| = \sqrt{\sum_{i=1}^{n} |Z_{0i} - Z_{1i}|^2}, \tag{3.3}$$

for two time series Z_0 and Z_1 of length n.

CORRELATION DISTANCE

Another well known distance measure is correlation distance, defined as one minus the Pearson correlation coefficient (see for example Deza and Deza, 2009). It is a suitable measure of similarity between two time series, when correlation in the data is expected (see Iglesias and Kastner, 2013). Correlation between two time series Z_0 and Z_1 is defined as:

$$\text{Cor}(Z_0, Z_1) = 1 - \frac{(Z_0 - \bar{Z}_0) \cdot (Z_1 - \bar{Z}_1)}{||Z_0 - \bar{Z}_0|| \cdot ||Z_1 - \bar{Z}_1||}, \tag{3.4}$$

with $\bar{Z}$ being the mean value of time series Z and $||\cdot||$ the Euclidean 2-norm as in Equation (3.3). We have to note here, that correlation cannot compare simple constant time series (leads to division by zeros) and is therefore not a distance metric in the sense of the definition Deza and Deza, 2009.

COMPARISON

For a comparison of the two distances for some example time series, see Figure 3.4. The example shows that the distance between two time series is not intuitively clear. The use of different distance metrics results in different sorting of distances between the shown pairs of time series. When normalising all time series (subtracting the mean and scaling by the standard deviation) correlation distance and Euclidean distance are equivalent (as shown for example by Deza and Deza, 2009). However, this leads to issues, when comparing to a constant time series (with zero standard deviation).

Both Euclidean distance and correlation are not taking into account the order of the values within each time series. For example, two identical time series that are shifted in time are seen as 'similar' with the correlation distance, but not as similar with the Euclidean distance and would not be considered as identical by either of them (see Figure 3.4). Additionally neither of the two distance metrics can deal with time series of different lengths or containing gaps.

3.3.3. CLUSTERING METHODS

Clustering methods for time series can be divided into two categories: feature based and raw data based (see for example Liao, 2005). Feature based methods first extract relevant features to reduce dimensionality (for example using Fourier- or wavelet-transforms) and then form clusters based on these features. They could also be used to deal with gaps in time series. We focus on the raw data based approach to not define features in advance and to make sure that no information within the data set is lost. We use three different

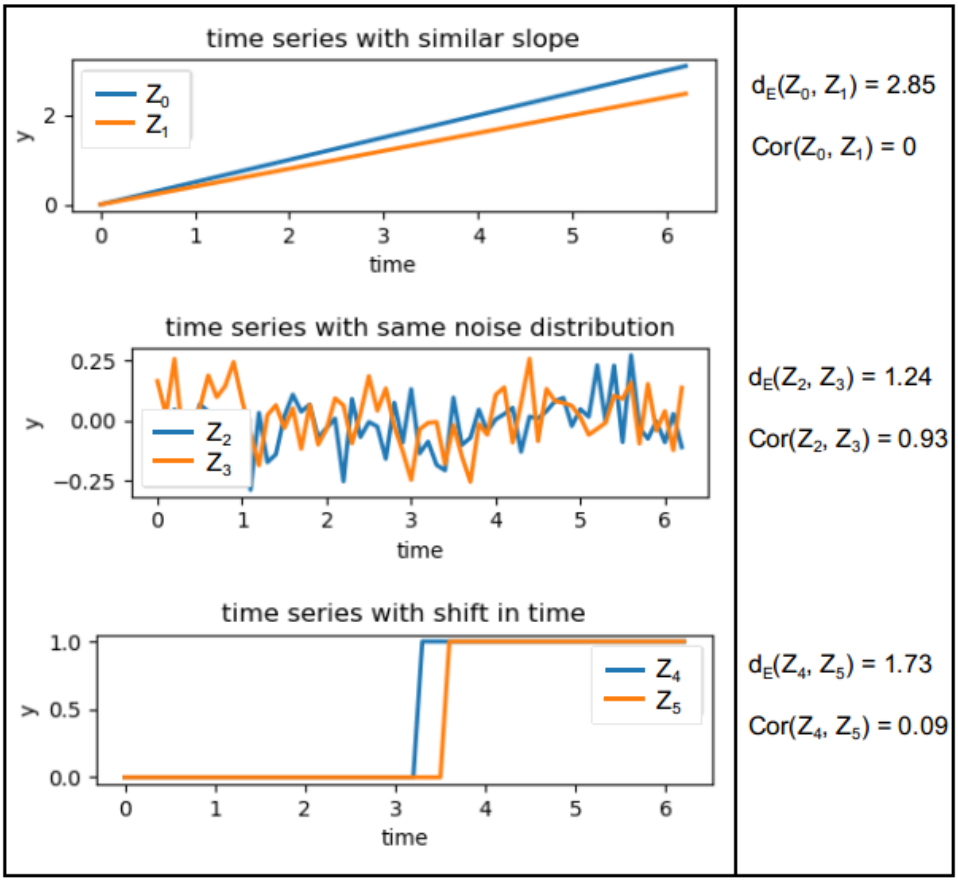

Figure 3.4: Example of three pairs of time series that are 'similar' to each other in different ways. The Euclidean distance would sort the differences as follows $d_E(Z_2, Z_3) < d_E(Z_4, Z_5) < d_E(Z_0, Z_1)$, whereas according to the correlation distance the order would be $\text{Cor}(Z_0, Z_1) < \text{Cor}(Z_4, Z_5) < \text{Cor}(Z_2, Z_3)$.

methods: k-means clustering, agglomerative clustering and Density-Based Spatial Clustering of Applications with Noise (DBSCAN). In Figure 3.5 an illustration of a partitioning of a simple 2D data set is shown for each of the three algorithms. The two clusters that can be distinguished in this example have different variances and are grouped differently by each of the algorithms.

For the implementation of all three algorithms, we make use of the Scikit-learn package in Python (see Pedregosa *et al.*, 2011).

K-MEANS CLUSTERING

The k-means algorithm was first introduced in 1955 and is still one of the most widely used clustering methods (Jain, 2010). The algorithm is based on minimising the sum J of

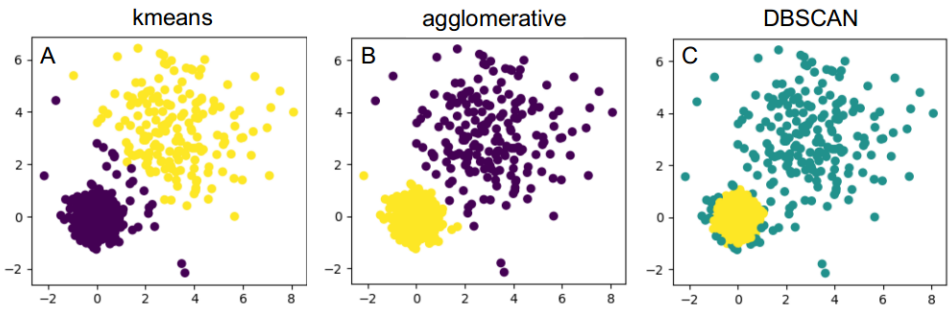

Figure 3.5: Example of clustering of data with two clusters with different variance: The k-means algorithm separates them, but adds a few points in the middle to the purple cluster instead of the yellow one (A). Agglomerative clustering separates both clusters according to their variances (B) and DBSCAN detects the cluster with low variance and high point density (yellow) and discards all other points as outliers (turquoise) (C).

all distances between points and centroids over all possible choices of k cluster centroids $V = \{v_1, \ldots, v_k\}$:

$$\mathrm{Min}_V J(V) = \sum_{j=1}^{k} \sum_{x_i \in v_j} ||x_i - v_j||^2, \tag{3.5}$$

with Euclidean distance metric $|| \cdot ||$. After the initial choice of k centroids among all points the following steps are repeated iteratively, until the above sum does not change significantly:

1. Assign each point to the cluster with closest centroid
2. Move centroid to mean of each cluster
3. Calculate sum of distances over all clusters (Equation (3.5))

Note that minimising the squared sum of distances over all clusters, coincides with minimising the squared sum of all within cluster variances. The convergence to a local minimum can be shown for the use of Euclidean distance (see for example Jain, 2010). The convergence is sped up using so-called k-means++ initialisation: After the random selection of the first centroid, all following centroids are chosen based on a probability distribution proportional to their squared distance to the already defined centroids. In this way the initial centroids are spread out throughout the data set and the dependence on the random initialisation of the cluster centroids is reduced.

There are variations of k-means using alternative distance metrics such as the L^1-norm (k-medoids, Park and Jun, 2009), however the convergence is not always ensured in these cases. Another issue to take into account when considering alternative distance metrics, is the definition of the cluster centroids as mean of time series, which is not automatically defined for any distance metric. For more information on k-means see Jain, 2010, Liao, 2005 and the documentation of the Scikit-learn package (Pedregosa *et al.*, 2011).

AGGLOMERATIVE CLUSTERING

Agglomerative clustering is one form of hierarchical clustering: It starts with each point in a separate cluster and iteratively merges clusters together until a certain stopping criterion is met. There are different variations of agglomerative clustering using different input parameter and stopping criteria (see for example Liao, 2005 or the documentation of the scikit-learn package (Pedregosa *et al.*, 2011)). We choose the minimisation of the sum of the within cluster variances using the Euclidean distance metric (Equation (3.5), where the centroids v_j are the mean values of the clusters) for a pre-defined number of clusters k. The algorithm starts with each point in a separate cluster and iteratively repeats the following steps until k clusters are found:

1. Loop through all combinations of clusters:
   - Form new clusters by merging two neighbouring clusters into one
   - Calculate squared sum of distances (Equation (3.5)) for each combination
2. Keep clusters with minimal squared sum of distances

In this way we use agglomerative clustering with a similar approach to the k-means algorithm, the same optimisation criterion with the same input parameter and Euclidean distance measure. We therefore expect similar results. However, this agglomerative clustering can easily be adapted to alternative distance measures and could therefore potentially deal with time series of different lengths or containing gaps.

DBSCAN ALGORITHM

Density-Based Spatial Clustering of Applications with Noise, DBSCAN, is a classical example of clustering based on the maximal allowed distance to neighbouring points that automatically derives the numbers of clusters from the data. It was introduced in 1996 by Ester *et al.*, 1996 and recently revisited by E. Schubert *et al.*, 2017. The algorithm is based on dividing all points into *core points* or *non-core points* that are close to core points but not themselves surrounded by enough points to be counted as core points. The algorithm needs the maximum allowed distance between points within a cluster (ε) and the minimum number of points per cluster (N_{min}) as input parameters. These two parameters define a core point: If a point has a neighbourhood of N_{min} points at ε distance, it is considered a core point. The algorithm consists of the following steps (E. Schubert *et al.*, 2017):

1. Determine neighbourhood of each point and identify core points
2. Form clusters out of all neighbouring core points
3. Loop through all non-core points and add to cluster of neighbouring core point if within maximal distance, otherwise classify as noise

In this way clusters are formed that truly represent a dense collection of 'similar' points. Since we choose to use correlation as distance metric, each cluster will contain correlated time series in our case. All points that can not be assigned to a close surrounding of a core point, are classified as noise or outliers.

3.3.4. EVALUATION CRITERIA

To determine if an algorithm is suitable, we expect that it fulfils the previously defined criteria:

- A majority of the observation area is separated into distinct regions,
- each cluster shows a change pattern that can be associated with a geomorphic deformation process, and
- time series contained in each cluster roughly follow the mean change pattern.

In order to establish these criteria, we compare the three clustering algorithms, as well as two choices for the number of clusters k, using the following evaluation methods.

VISUAL EVALUATION

The clustered data are visualised in a top view of the observation area, where each point represents the location of a grid cell. Each cluster is associated with its cluster centroid, the mean elevation time series of all time series in the respective cluster. For visualisation purposes we have added the median elevation back to the cluster centroids, even though it is not taken into account during the clustering. We subsequently derive change processes visually from the entire clustered area. We establish which kind of deformation patterns can be distinguished and estimate rates of change in elevation and link them to the underlying process.

QUANTITATIVE EVALUATION

We use the following criteria to compare the respective clustering and grid generation methods quantitatively:

- percentage of entire area clustered
- minimum and maximum within cluster variation
- percentage of correctly identified change in test area with bulldozer work

The percentage of the area that is clustered differs depending on the algorithm. Especially DBSCAN sorts out points that are too far away (i.e. too dissimilar) from others as noise. This will be measured over the entire observation area. The number of all complete time series counts as 100%.

Each cluster has a mean centroid time series and all other time series deviate from that to a certain degree. We calculate the average standard deviation over the entire month per cluster and report on the minimum and maximum value out of all realised clusters.

TEST AREA

To allow for a comparison of the clusters with a sort of ground truth, we selected a test area at the bottom of the footpath. In this area a pile of sand was accumulated by a bulldozer, after the entrance to the path was covered with lots of sand during a period of rough weather conditions (8 to 16 January, corresponding to day 7 to 14 in our time

series), as reported by Anders *et al.*, 2019. We chose two time stamps for illustration, and show the elevation before the bulldozer activity at the end of the stormy period on 16 January, after the bulldozer activity on 18 January and the difference between the elevations on these two days in Figure 3.6 (first row, A,B,C). The area does not change significantly after this event. Within this test area we classify (manually) each point as 'stable' or 'with significant change' depending on a change in elevation of more than 5 cm (positive or negative). Then we evaluate for each clustering method if the points that are classified as 'with significant change' are in a separate cluster than the 'stable' points.

The stable cluster consists of cluster 0, the largest cluster when using $k = 6$ for k-means and agglomerative clustering and cluster 0 and 1 combined in the case of $k = 10$ clusters. For evaluating the results of the DBSCAN algorithm we consider all locations that are not clustered (noise) and points in cluster 1 as the 'stable' areas, because the average erosion in cluster 1 is less than 0.15 cm per day. We do not distinguish if there are different clusters within the category of 'with significant change'. However, in Figure 3.6, the different clusters can be distinguished by their colours, corresponding to the colours of the clusters shown in subsequent figures (Figures 3.7, 3.8 and 3.9). We then compare the percentage of correctly classified grid points for the test area, for each of the grid generation and clustering methods.

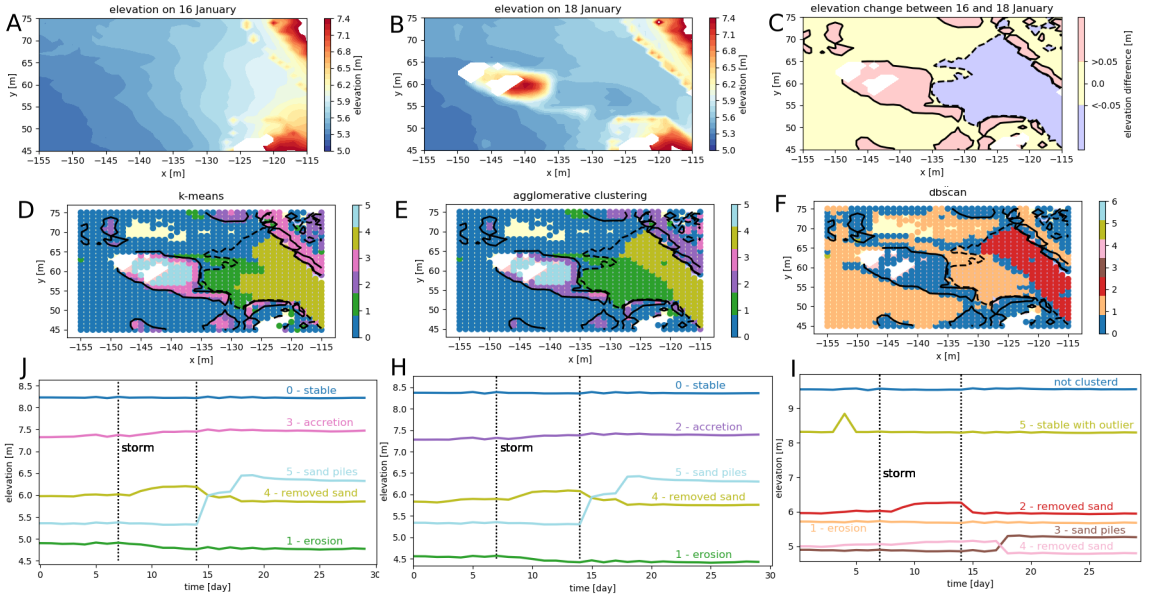

Figure 3.6: Test area for the comparison of clusters generated with three different algorithms. The test area is located where the northern access path meets the beach (see Figure 3.1). 1st row: The elevation in the test area is shown on the day before the bulldozer accumulated a sand pile, when the entrance of the path was covered in sand (A) and after the bulldozer did its job (B). Behind the sand pile appears a gap in the data (in white), as the sand pile is obstructing the view for the laser scanner. To the right we show the difference in elevation between January 16 and 18 from a significant level upwards (red) and downwards (blue) (C). 2nd row: Test area with significant changes in elevation (contour lines) and points clustered using the k-means algorithm (D), agglomerative clustering (E) and the DBSCAN algorithm (F). The colours of the clustered dots represent the clusters as shown in Figures 3.7, 3.8 and 3.9, respectively. 3rd row: The corresponding mean time series for each of the relevant clusters are displayed below each of the plots (G,H,I). The dotted lines mark the beginning and end of a stormy period.

3.4. RESULTS

The results are presented in two parts. First, we compare two different choices of the parameter k for the k-means algorithm and for agglomerative clustering. Then, we compare all three clustering methods and evaluate results on the test area, where a bulldozer created a pile of sand (as indicated in Figure 3.1) and in terms of percentage of data clustered, standard error within each cluster and physical interpretation of clusters.

3.4.1. CLUSTERING

For the k-means algorithm and agglomerative clustering, we consider two different values $k = 6$ and $k = 10$, exemplary for a smaller number of clusters and a higher number of clusters.

Table 3.2: Summary of comparison of k-means algorithm, agglomerative clustering (AGG) and DBSCAN algorithm.

	k-means		**AGG**		**DBSCAN**
entire observation area					
number of clusters	6	10	6	10	6
min no. points/cluster	108	34	108	39	45
area clustered	100%	100%	100%	100%	44%
max std error/cluster	3.22 m	3.1 m	3.18 m	2.86m	4.0 m
min std error/cluster	0.77 m	0.68 m	0.79 m	0.71 m	0.33 m
test area: correctly identified					
stable points	81%	82%	86%	86%	99%
positive changes	97%	97%	86 %	86%	0%
negative changes	93%	93%	98 %	98%	54%
total	85%	86%	88 %	88%	79%

K-MEANS

With the k-means algorithm, the entire observation area is partitioned. The resulting partition depends on the random initialisation. The standard error within each cluster is relatively high, compared to the stable area (see Table 3.1) and generally increases with the size of the cluster. Even the cluster with the smallest standard error (averaged standard deviation per time series over the clustered area), still shows a standard error of 0.77 m (cluster 5 for $k = 6$). We show the resulting clusters obtained using the k-means algorithm with number of clusters $k = 6$ and $k = 10$. Visual inspection shows that both values lead to good, usable results by partitioning the set of time series into clusters that are small enough to capture geomorphic changes but not too large, which would make them less informative. As displayed in Figure 3.7, a large part of the beach is contained in a 'stable' cluster when using $k = 6$ (cluster 0, blue). This cluster, as well as some of the others, are split up into several smaller clusters when using $k = 10$. For example, the intertidal zone (i.e. the area that is under water during high tide and exposed during low

tide) is eroding mostly during stormy days in the first half of the month. This zone is contained entirely in cluster 1 (green) when using $k = 6$. In the case of $k = 10$, this part is split up into three clusters, one with a similar mean time series (cluster 2, green), one eroding with a pattern similar to cluster 2, but mostly representing sand banks (cluster 3, brown) and one gradually eroding at a low rate over the entire month (cluster 1, orange). It also becomes clear, that the sand piles that were generated by bulldozer works at different locations ($k = 6$ cluster 5, light blue) were created on different days ($k = 10$, clusters 8 and 9, yellow and light blue). Some features, like the cleared part of the paths, the sand piles and the intertidal zone can be distinguished in both cases.

On the test area the k-means algorithm correctly classifies about 85% of points into 'stable', 'significant negative change', or 'significant positive change' in the case of $k = 6$. However, as can be seen in Figure 3.6, a part of the points with negative change are not identified. These clusters are split up further in the case of $k = 10$, which does not influence the results in the test area a lot. A summary of these results is provided in Table 3.2.

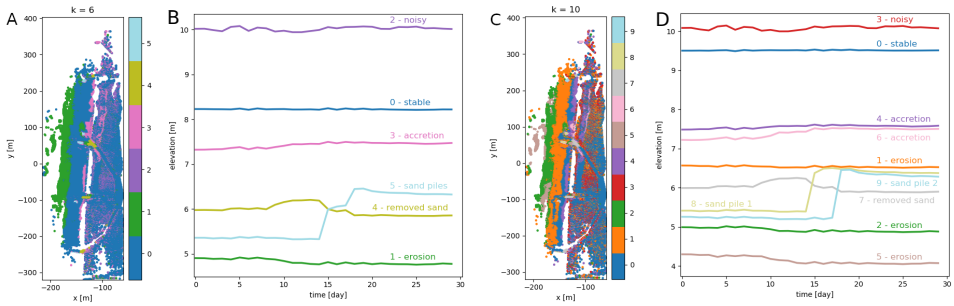

Figure 3.7: A/C: Overview of the entire observation area divided into clusters using k-means with $k = 6$ (A) and $k = 10$ (C). B/D: Corresponding cluster centroids for each of the clusters shown in A and C, respectively. By using a larger number of clusters k, more processes become visible, for example two sand piles (A/B: cluster 5) created on two different days (C/D: cluster 8 and 9). Also the large stable areas (A/B: cluster 0) and slowly accreting areas (A/B: cluster 3) are split up into several clusters: A slightly eroding area (C/D: cluster 3) is split up from the stable part and the accreting area is split into two (C/D: cluster 4 and cluster 6).

AGGLOMERATIVE CLUSTERING

The agglomerative clustering algorithm is set up, as the k-means algorithm, to find six and ten clusters. It produces results very similar to the clusters found with the k-means algorithm, as can be seen comparing Figures 3.7 and 3.8 and Figures 3.6 D and E. Clusters 2 and 3 from agglomerative clustering correspond roughly to the clusters 3 and 2 from k-means clustering. The ordering of clusters is according to size, so more time series are considered 'noisy' according to k-means, whereas agglomerative clustering assigns more of these time series to the gradually accreting cluster. All other clusters appear to be nearly identical and show similar spatial distributions as well as centroid shapes. The differences between the two choices of the number of clusters k are also very similar.

On the test area, the detection of negative and positive changes is more balanced and leads to an overall score of 88 % correctly identified points. Agglomerative clustering clearly separates the path that was cleared by the bulldozer and identifies it as eroding.

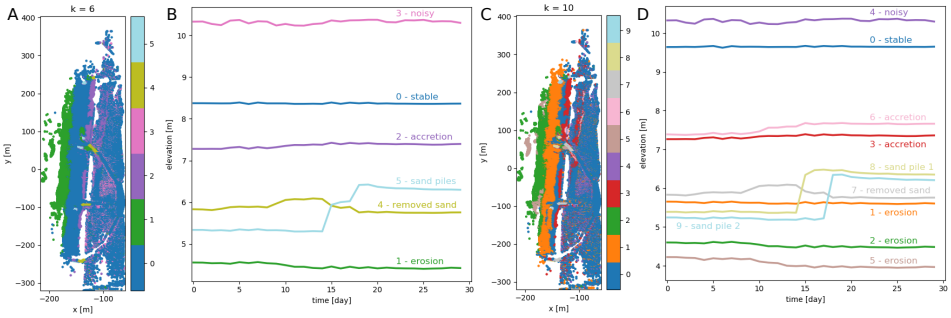

Figure 3.8: A/C: Overview of the entire observation area divided into clusters using agglomerative clustering with $k = 6$ (A) and $k = 10$ (B). B/D: corresponding cluster centroids for each of the clusters shown in A and C respectively. The clusters are similar to the ones found with k-means.

DBSCAN

When we use the DBSCAN algorithm on the same data set, with minimum number of points $N_{min} = 30$ and maximum distance $\varepsilon = 0.05$, a large part of the time series (55 %) is classified as noise, meaning that they are not very similar (i.e. not correlated, since we use correlation as distance measure) to any of the other time series. However they roughly match the combined areas that are identified as stable and noisy by k-means (clusters 0 and 2 for $k = 6$). The remaining time series are clustered into six clusters. The standard error within each cluster is generally lower than in the clusters generated with k-means (minimum standard error is 0.33 m) without considering the time series that are classified as noise.

The intertidal zone cannot be separated clearly from the 'noise' part of the observation area, nor can we distinguish the stable path area or the upper part of the beach. In the test area, the sand pile is not represented by a separate cluster and positive changes in elevation are not found, which results in an overall worse percentage of correctly identi-

fied points. But, two clusters represent areas, which are relatively stable throughout the month, except for a sudden peak in elevation on one day. These peaks are dominated by a van parking on the path on top of the dunes and people passing by, and are not caused by actual deformation, compare Figure 3.9.

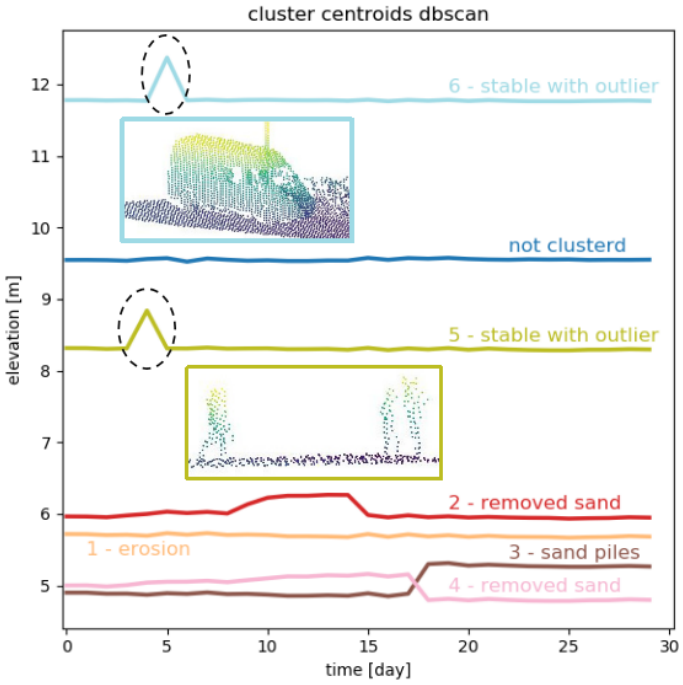

Figure 3.9: Mean time series per cluster found with the DBSCAN algorithm. Outliers or not clustered points are represented by the blue mean time series. The two most prominent time series (cluster 5 and 6, light green and light blue) are located on the path on top of the dunes. The peaks are caused by a group of people and a van, on the 5th and 6th of January respectively, illustrated by the point clouds in the middle of the plot.

On the test area the DBSCAN algorithm performs worse than both other algorithms. In total 79% of points are correctly classified into 'stable' or 'significant negative change'. As stable points we count in this case all points that are classified either as noise or belong to cluster 1 (orange). The reason for this is that the mean of all time series that are not clustered appears relatively stable, while cluster 1 describes very slow erosion of less than 0.15 cm per day. This matches with 99% of points classified as stable in the ground truth data. But, no single cluster is formed containing only the points where sand is accumulating, even though these clusters are distinguished by the other two algorithms. These points are mixed up with the large cluster of slightly eroding points in cluster 1. We can see in Figure 3.6, that the only significant process found in the test area is the

cleared path (cluster 2, red).

3.4.2. IDENTIFICATION OF CHANGE PROCESSES

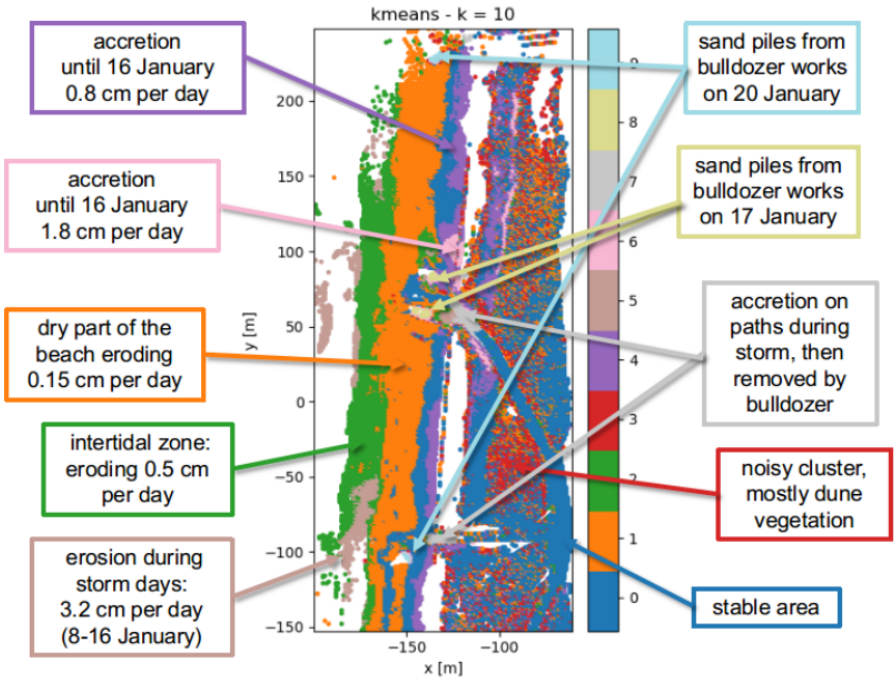

Figure 3.10: Observation area partitioned into clusters by the k-means algorithm with $k = 10$. The associated processes are annotated with the corresponding colours.

Considering the clusters found by the k-means algorithm and agglomerative clustering, we can clearly distinguish between time series that represent erosion and accretion with different magnitudes and at different times of the month, as well as a sudden jump in elevation, caused by bulldozer work. In Figure 3.10 we show the clusters and associated main process. To give an idea of the magnitude of the most prominent change patterns, we fit straight lines through the mean time series or parts of it (where the slope is steepest) and derived average rates of change in elevation from the estimated slopes. The clusters dominated by erosion, close to the water line (clusters 2 and 5) represent the inter-tidal zone of the beach. The elevation changes in this area are likely caused by the effects of tides and waves. The change rates were partly accelerated during the stormy period in the first half of the month. Accreting areas are mostly at the upper beach, close to the dune foot and on the paths in the dunes (clusters 4, 6 and 7). These areas as well as a large cluster on the upper beach (cluster 1, orange), which undergoes a slight and gradual erosion over the entire month, are likely dominated by aeolian sand transport. The most obvious change process is the sand removed from the entrances of the paths leading to the beach by bulldozer works (cluster 7) and accumulated in piles of sand at four different locations on two days (clusters 8 and 9). Points contained in the noisy cluster (cluster 3) are spread out through the dune area and noise is probably caused by moving vegetation.

3.5. DISCUSSION

We successfully applied the presented methods on a data set from PLS and demonstrated the identification of deformation processes from the resulting clusters. Here we discuss our results on distance measures, clustering methods and the choice of their respective input parameters and derivation of change processes.

3.5.1. DISTANCE MEASURES

Possible distance measures for the use in time series clustering are analysed among others by Iglesias and Kastner, 2013 and Liao, 2005. We use Euclidean distance in combination with the k-means algorithm and agglomerative clustering for our analysis. It has been shown by Keogh and Kasetty, 2003 that especially for time series with high dimensions, alternative distance measures rarely outperform Euclidean distance. However, we have to note here, that Euclidean distance is affected by the so called 'curse of dimensionality', which causes a space of long time series (with many dimensions) to be difficult to cluster. The problem with clustering time series in high dimensional spaces with k-means, is that Euclidean distance is based on the sum of all point wise differences. This leads to a space, where the variance of the distances decreases with increasing time series length. Therefore it will be harder to categorise time series as similar, and fewer meaningful clusters will emerge, the more observations we use. This could possibly lead to difficulties, when extending these methods to the use of longer time series, but does not appear to degrade results on our current data set. For more details on this issue see Assent, 2012, Verleysen and François, 2005 and Zimek *et al.*, 2012.

We chose for the use of correlation distance with the DBSCAN algorithm, because correlation in principle represents a more intuitive way of comparing time series (see Fig-

ure 3.4). DBSCAN is based on identification of clusters of high density, which in our case works better using correlation distance instead of Euclidean distance. Using Euclidean distance, there are very few clusters of 'very similar' time series and an even larger part of the beach is classified as noise. Only in combination with correlation distance, we could derive a set of input parameters for the DBSCAN algorithm to produce relevant results.

Scaling the time series with their respective standard deviations for the use of Euclidean distance would make these two distance measures equivalent. However, this did not improve our results using k-means or agglomerative clustering. Subtle differences within the stable cluster would become prominent in that case, but the larger differences between clusters as we find them without the scaling, would be reduced.

Neither of the two distance measures analysed here can deal with gaps in the time series, which would be of great interest to further analyse especially the intertidal area and sand banks. Additionally, both distance measures do not allow to identify identical elevation patterns that are shifted in time as similar. An alternative distance measure suitable to deal with these issues would be Dynamic Time Warping (DTW), which accounts for similarity in patterns even though they might be shifted in time (Keogh & Ratanamahatana, 2005). An interpolation method to fill gaps in elevation over short time spans based on surrounding data or a feature based clustering method could be other alternatives.

3.5.2. CLUSTERING METHODS

The use of k-means clustering on elevation time series from the same data set was demonstrated by Lindenbergh *et al.*, 2019 and has shown promising first results. We follow the same approach and, as a comparison, use agglomerative clustering, with the same optimisation criterion, distance metric and input parameter. As expected the results are similar, although agglomerative clustering does not depend on random initialisation. It therefore delivers the same result for every run, which is an advantage. Considering our previously defined criteria:

- a majority of the observation area is separated into distinct regions,
- each cluster shows a change pattern that can be associated with a geomorphic deformation process, and
- time series contained in each cluster roughly follow the mean change pattern,

both algorithms are suitable and the differences in the resulting clusters are negligible for our specific data set.

However, the computational effort needed to loop through all possible combinations of merging clusters for agglomerative clustering is considerably higher. Of the three algorithms that were used in this study, agglomerative clustering is the only one that regularly ran into memory errors. This is a disadvantage considering the possible extension of our method to a data set with longer time series.

One of the disadvantages of the k-means algorithm and our configuration of agglomerative clustering, is that the number of clusters has to be defined in advance. Our choices of $k = 6$ and $k = 10$ clusters both yield promising results, but remain somewhat arbitrary,

especially without prior knowledge of the data set. A lower number of clusters k (for example $k = 6$) yields a division of the beach into sections (inter-tidal zone, dry part of the beach) and highlights the most prominently occurring changes (bulldozer works). When using a larger number of clusters k, several of the previously mentioned clusters are split up again and more detailed processes become visible. The erosion and accretion patterns on the beach appear at different degrees in distinct regions, which is valuable information. Also the sand piles, which appeared in one cluster for $k = 6$ are now split up according to the different days, on which they were generated. We consider this possibility to identify and specify anthropogenic induced change an illustrative example of the influence of the choice of the number of clusters k. We have considered two data independent methods to determine a suitable value for k: analysis of the overall sum of variances for different values of k and so-called *cluster balance* following the approach of Jung *et al.*, 2003. Neither of them resolved the problem satisfactorily and we cannot make a generalised recommendation, independent of the application, for the choice of k at this point.

To avoid this issue we also compare both approaches with the use of the DBSCAN algorithm. It is especially suitable to distinguish anomalies and unexpected patterns in data as demonstrated by Çelik *et al.*, 2011 using temperature time series. To decide, which values are most suitable for the two input parameters of the DBSCAN algorithms we plot the percentage of clustered points and the number of clusters depending on both parameters (see Figure 3.11). However, this did not lead to a clear indication of an 'optimal' set of parameters. After the trade-off analysis between the number of points in clusters and the number of clusters (not too high, so that the clusters become very small and not too low so that we generate only one big cluster) we chose $\varepsilon = 0.05$ and $N_{min} = 30$ by visually inspecting the resulting clusters.

An alternative clustering approach for time series based on fuzzy C-means is proposed by Coppi *et al.*, 2010. They develop a method to balance the clustering based on the pattern of time series while keeping an approximate spatial homogeneity of the clusters. This approach was successfully applied to time series from socio-economic indicators and could be adapted for our purpose. It could potentially improve detection of features like sand bars, or bulldozer work, but not distinguish moving vegetation in the dunes as our current approach does.

A similar approach would be to use our clustering results and identified change patterns as input to the region-growing approach of Anders *et al.*, 2020. In this way we could combine advantages of both methods by making the identification of the corresponding regions for each distinct deformation pattern more exact, without having to define possible deformation patterns in advance.

3.5.3. DERIVATION OF CHANGE PROCESSES

As shown in Figure 3.10, we identified change processes from the clusters generated by k-means. Agglomerative clustering shows similar clusters and therefore yields similar results. Each centroid representing the mean time series of the k-means clusters shows a distinct change pattern (see Figures 3.7 and 3.8), which allows to conclude on a predominant deformation process. By fitting a straight line through the mean time series, or part of it, we estimated the slope corresponding to the average rate of change in elevation.

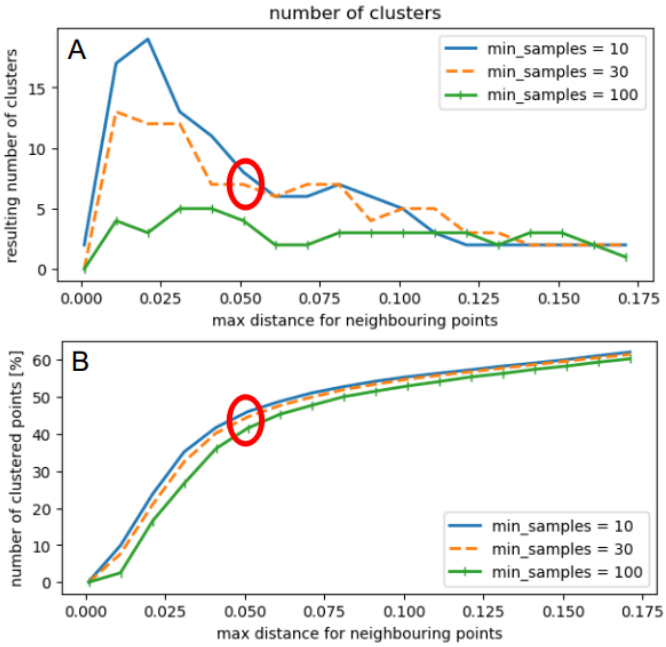

Figure 3.11: DBSCAN selection of input parameters: Number of clusters versus input parameter maximum distance within clusters and minimum number of points and percentage of total points in clusters (not classified as noise/outliers). The choice of an 'optimal' set of parameters is not obvious. We indicate our selection with a red circle in both plots.

Associating the centroids with the location and spatial spread of the clusters allows to derive the main cause for the respective deformations. In some cases extra information, or an external source of validation data would be useful to verify the origin of the process. This will be taken into account for future studies. The location of the clusters and for example the steep rise of the mean time series representing the sand piles allows for the conclusion that the cause of this sudden accretion is anthropogenic. The information found by Anders *et al.*, 2019 for the research on their study confirms the coinciding bulldozer works. The derived average rates of change in elevation allow for the possibility to derive mass budgets to quantify volume changes over specific amounts of time from our data, showing a possible application of our method, that is of large scientific interest (see for example Schipper *et al.*, 2016).

The DBSCAN algorithm successfully identifies parts of the beach that are dominated by a prominent peak in the time series (caused by a van and a small group of people). Out of the three algorithms that we compare, it is most sensitive to these outliers in the form of people or temporary objects in the data. It was not our goal for this study, to detect people or objects on the beach, but this ability could be a useful application of the DBSCAN algorithm to filter the data for outliers in a pre-processing step.

3.6. CONCLUSIONS

We compared three different clustering algorithms (k-means, agglomerative clustering and DBSCAN) on a subset of a large time series data set from PLS on a sandy urban beach. We successfully separated the observed beach and dune area according to their deformation patterns. Each cluster, described by the mean time series, is associated with a specific process (such as bulldozer work, tidal erosion) or surface property (for example moving vegetation cover).

The most promising results are found using k-means and agglomerative clustering, which both correctly classify between 85 and 88 % of time series in our test area. However, they both need the input of the number of clusters we are looking for and agglomerative clustering is computationally expensive. DBSCAN turned out to be more suitable for the identification of outliers or unnatural occurring changes in elevation due to temporary objects or people in the observed area.

Our key findings are summarised as follows:

1. Both k-means and agglomerative clustering fulfil our criteria for a suitable method to cluster time series from PLS.

2. Predominant deformation patterns of sandy beaches are detected automatically and without prior knowledge using these methods. The level of detail of the detected deformation processes is enhanced with increasing number of clusters k.

3. Change processes on sandy beaches, which are associated with a specific region and time span, are detected in a spatio-temporal data set from PLS with the presented methods.

Our results demonstrate a successful method to mine a spatio-temporal data set from PLS for predominant change patterns. The method is suitable for the application in an automated processing chain to derive deformation patterns and regions of interest from a large spatio-temporal data set. It allows such a data set to be partitioned in space and time according to specific research questions into phenomena, such as for example the interaction of human activities and natural sand transport during storms, recovery periods after a storm event or the formation of sand banks. The presented methods enable the use of an extensive time series data set from PLS to support the research on long-term and small scale processes on sandy beaches and improve analysis and modelling of these processes. In this way we expect to contribute to an improved understanding and managing of these vulnerable coastal areas.

4

IDENTIFYING TOPOGRAPHIC CHANGES AT THE BEACH USING MULTIPLE YEARS OF PERMANENT LASER SCANNING

Mieke KUSCHNERUS, Sander VOS, Sierd DE VRIES, José A. Á. ANTOLINEZ, Roderik LINDENBERGH,

Sandy beach-dune systems make up a large part of coastal areas world wide. Their function as an eco-system as well as a protective barrier for human and natural habitat is under increased threat due to climate change. A thorough understanding of change processes at the sediment surface is essential to facilitate prediction of future development and management strategies to maintain their function. Especially slow and small scale processes happening over several days up to weeks at cm level, such as aeolian sand transport are difficult to identify and analyse. Permanent laser scanning (PLS) is a useful tool in the study and analysis of coastal processes as it captures a data representation of the evolution of the sediment surface over extended periods of time (up to several years) with high detail (at cm-dm level). The PLS data set considered for this study, consists of hourly acquired 3D point clouds representing the surface evolution of a section of the Dutch coast during three years. However, it is challenging to extract concrete information on specific change processes from the large and complex PLS data set. We use multiple hypothesis testing in order to reduce the PLS data set to a so-called inventory of trends, consisting of 12.8 million partial time series with associated rate of change and elevation. The inventory of trends proofs to be a suitable tool to identify natural processes such as storms and

This chapter has been published in Coastal Engineering **193**, 104594 (2024) Kuschnerus, de Vries, *et al.*, 2024.

aeolian sand transport in our test area in the aeolian zone of a sandy beach-dune system on the Dutch coast. We identify these processes and provide a tool to derive summarising data from the complex PLS data set. We find that all partial time series identified as most likely representing aeolian sand transport, result in 1354 m^3 of sand deposition in our study area over the course of three years. We also show a comparison with transects from JarKus data and find a correlation between anthropogenic activities and erosion in our test area with a correlation coefficient of 0.3.

4.1. INTRODUCTION

Coastal areas worldwide serve a multitude of purposes. Specifically sandy beach-dune systems (SBDS), which make up 31 % of non-iced coast (A. Luijendijk *et al.*, 2018), serve as a protective barrier to prevent flooding, as habitat for a large number of species, and as recreational areas. The resilience of SBDS is under increasing pressure due to both natural and human driven climate change (Grabemann & Weisse, 2008; Intergovernmental Panel on Climate Change (IPCC), 2014). While sea level rise and increased severity of storm events are foreseen, the growing human population continues to develop infrastructure on and near coastal areas worldwide, which under improvisational planning can lead to coastal squeeze (Vousdoukas *et al.*, 2020). To mitigate these effects, predictions of the development of SBDS are needed, which require a thorough understanding of topographic changes at the sediment surface of SBDS.

Topographic coastal changes at SBDS are driven by anthropogenic, ecologic and geomorphological processes at different spatial and temporal scales. These driving influences also interact. While a storm can cause significant erosion of the beach-dune profile over the course of several hours (van Wiechen *et al.*, 2023), wind driven sediment transports may only cause significant accretion of the dune profile of comparable extent over the course of several decades (de Vries *et al.*, 2012). In addition, the net result of the interaction between erosive and accretive processes is sensitive to even small changes in any of the influencing factors (Stive *et al.*, 2002). Thus, human interventions, like shore nourishment, frequent beach maintenance or buildings on the beach might influence these as well (Poppema *et al.*, 2021; Schipper *et al.*, 2016). The knowledge needed to analyse past and present development in order to make viable predictions of the future, can be gained by observing, quantifying and subsequently interpreting these dynamics and interactions at the sediment surface at different time scales.

Given the difference in scale, processes and interactions may be challenging to identity and analyse through sparse observations of surface beach dynamics. Regular and frequent observations at high level of detail are desired to cover all temporal and spatial scales. Optical as well as radar remote sensing are suitable observation tools, as well as in-situ measurements with sensors and through inspections.

Yearly observation of the Dutch coast with airborne laser scanning provide a tool to monitor the long term development of surface dynamics in the Dutch beach-dune system. The resulting JarKus transects (Rijkswaterstaat - Dutch Ministery of Infrastructure and Water Management, 2022) were for example analysed by C. O. IJzendoorn *et al.*, 2021. These transects, regularly spaced at 250 m, are used to study the location of the coastline (quantified as momentary coastline (MKL)) and for decision making in coastal management in the Netherlands.

Several other studies use incidental terrestrial laser scanning to observe the effect of one specific event or process (see for example Ackerley *et al.*, 2016; J. E. Schubert *et al.*, 2015). PLS is a tool to continuously monitor a section of coast over a longer period at high temporal frequency (up to hourly) and small spatial resolution (up to cm) (see Anders *et al.*, 2019; Vos *et al.*, 2017). PLS has been used in combination with analysis of video images to detect bulldozer activities on the beach (Barbero-García *et al.*, 2023) and for several other applications, for example glacier observation (Voordendag *et al.*, 2021) and rock fall monitoring (Schröder *et al.*, 2022).

Anders *et al.*, 2020 developed the so-called 4D-objects by change to find spatially connected areas in which time series behave in a similar way. Each 4D object by change can then be associated with a physical process, that would typically cause this pattern. Their method favours prominent changes and only finds affected areas that are connected, like for example bulldozer works or the appearance of a sand bar. A method to group and classify the 4D-objects by change was developed by Hulskemper *et al.*, 2022. In our previous study on a similar data set of shorter length (covering six months), we used clustering algorithms to group time series with similar change behaviour to identify areas affected by the same physical processes (Kuschnerus, Lindenbergh, & Vos, 2021). This method turned out to be useful for finding areas affected by bulldozer works as well as different erosion rates in the intertidal area. However, it is less suitable to process our entire three-year data set. Winiwarter *et al.*, 2023 used Kalman filtering to process long time series containing gaps and found those time series to be a suitable input to feature based clustering methods.

The aim of this paper is to identify topographic changes at the sediment surface of the beach-dune system using PLS on a section of the Dutch coast. The permanently installed instrument measures with high frequency (hourly) and covers a section of about 800 m length in high detail (cm to dm level of elevation change). The acquired data set captures the surface dynamics but its structure of 3D point clouds is cumbersome and information is not easily extracted. We propose a method to extract elevation time series and their rates of change and collect them as an *inventory of trends* in order to reduce and simplify the data set. The trend inventory allows to identify and analyse observed dynamics over a period of three years. We demonstrate the link but also the additional value compared to the JarKus data set by relating the observations in the respective matching locations. Specifically we aim to answer the following research questions:

1. What elevation trends at the sediment surface can be derived from the PLS data set?

2. How do high resolution PLS data compare to yearly JarKus-data?

3. How can small and slow changes at the sediment surface be identified and quantified with PLS?

4. How can the effect of anthropogenic activities on the dynamics at the sediment surface be analysed using PLS data?

The novelty of the proposed work is, firstly, the presented method to generate a simpler representation of our data set for efficient data mining. Secondly, we process for the first time the entire hourly three-year data record from PLS, which has proven to be challenging with previously developed methods. Finally, our methods allow to identify surface dynamics in the aeolian zone, which are difficult to study with other types of data and methods (see for instance Uphues *et al.*, 2022). Especially our new method is capable of capturing long term but small-scale changes due to aeolian sand transport.

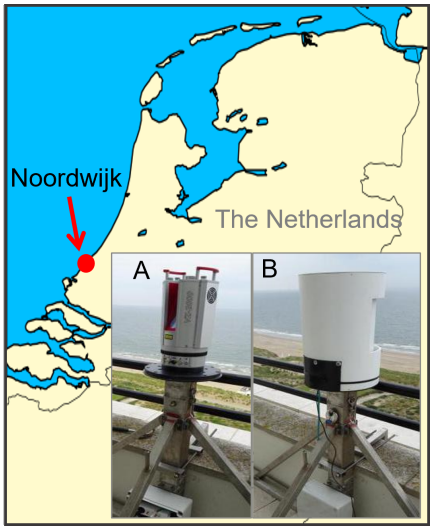

Figure 4.1: The study site on the Dutch coast in Noordwijk, The Netherlands, with the laser scanner mounted on a hotel balcony without (A) and with (B) protective cover.

4.2. FIELD SITE AND DATA SET

The study area is part of a typical example of a Dutch urban beach and provides the infrastructure needed for the instrumental set-up. The Dutch beach-dune system is orientated from south-southwest to north-northeast and consists of quarts sand (C. IJzendoorn *et al.*, 2023). The coast is driven by meso-tidal tides of about 2 meters height (Wijnberg, 2002), waves of about 1.7 meter height during average conditions and 5 meter height during storm conditions (A. P. Luijendijk *et al.*, 2017) and with south westerly winds (Vos *et al.*, 2024). The Dutch coast is maintained regularly (Maij-Weggen, 1990) with (underwater) sand nourishments resulting in an accreting environment around Noordwijk (de Vries *et al.*, 2012). The observed area includes a sandy beach and dunes, covered with vegetation and is about 1 km long and 250 m wide. The beach is strongly influenced by tides and varies in width between 80 m and 140 m under normal weather conditions. In rare occasions, storm reduces the width of the beach even further. Figure 4.2 shows an overview of the study area with an aerial image and the top view of a point cloud as an example. In the point cloud the colours are based on elevation. We focus in this study on a 400 m long part of the aeolian zone and do not include the intertidal zone or dunes. Within the study area there is often a beach scarp of up to 1 m height, see Figure 4.2B, which can lead to sudden changes in elevation observations.

To identify and analyse change processes, we use an exemplary data set from permanent lasers scanning (PLS) consisting of point clouds collected hourly during a period of three years from July 2019 to June 2022. The entire data set is published and can be accessed through 4TU.ResearchData Vos *et al.*, 2023. The point clouds processed for this study are acquired with a Riegl VZ-2000 laser scanner, which is permanently mounted

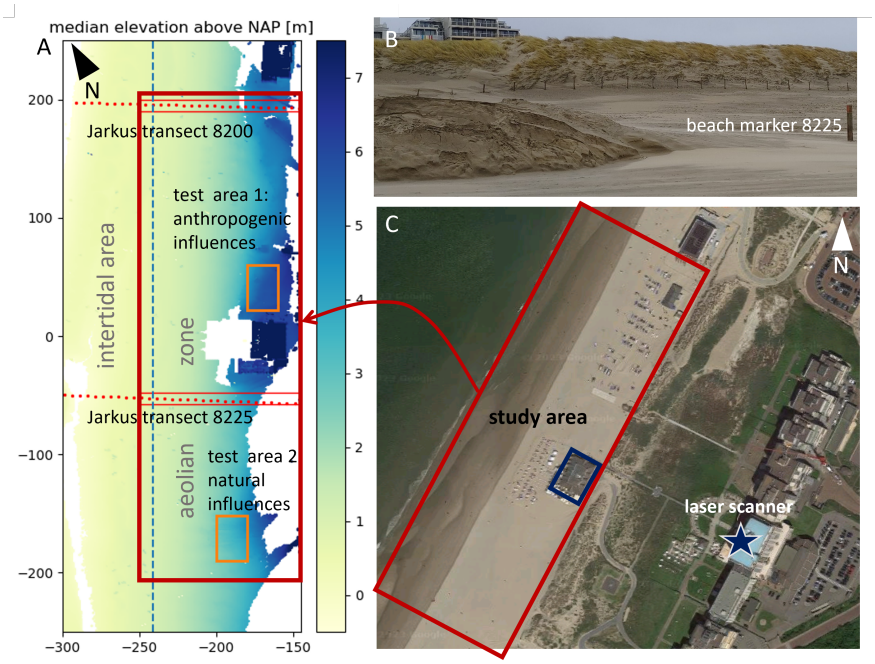

Figure 4.2: Overview of study area: Point cloud of the sandy beach (A), photo of a beach scarp formed at the edge of the aeolian zone (B) and aerial image of Grand Hotel Huis ter Duin with the adjacent dunes and beach (C, by Google maps). In the point cloud the colours are based on elevation. White areas appear because of missing data in the point clouds in the shadow zones of dunes and buildings, or because of water present in the intertidal zone. The test areas are marked with a square: test area 1 is next to a beach cafe and influenced by many anthropogenic changes (i.e. bulldozer works, frequent visitors) and test area 2 is mostly influenced by natural processes. In photo (B) the beach marker at location of Jarkus transect with ID 8225 is visible. The location of the laser scanner is marked with a star in the aerial image.

on the balcony of Grand Hotel Huis ter Duin in Noordwijk, The Netherlands. The laser scanner is mounted on a metal frame at 55.757 m height above NAP that is fixed to the balcony to maximise stability and covered with a protective cover (see Figure 4.1).

4.2.1. INSTRUMENT SETTINGS

The settings of the laser scanner are summarised in Table 4.1. Data is collected by running a scan of nearly 180 horizontal degrees covering a stretch of beach of just under 1 km long every hour with an angular spacing of 0.03 degree.

The laser scanner generates a 3D point cloud per scan, made up of x,y,z-coordinates which are calculated by the Riegl proprietary software out of recorded range, horizontal

range accuracy (at 150 m range) [m]	0.008
angular spacing [deg]	0.003
beam divergence [mrad]	0.3
inclination sensor frequency [Hz]	1
measurement accuracy [deg]	0.008

Table 4.1: Specifications of Riegl VZ-2000 laser scanner according to documentation.

total number of files	20 875
number of good quality files	19 386
days scanned	954
interruptions (> 24h)	21
number of points in study area	800 000
range	145 - 500 m
point density	1 - 40 pt/m^2
footprint size	0.02 - 0.1 m^2
inclination angle	70 - 90 deg

Table 4.2: Data properties considering all hourly scans between 11 July 2019 and 21 June 2022.

and vertical angle data, as well as intensity of the backscatter signal, per point. The internal inclination sensor records inclination values during each scan, with a frequency of 1 Hz (not matching the scanning frequency). A mean value and the standard deviation of these inclination angles is used for correction of tilts in the scanner (see section 4.3.1).

There are some temporal gaps in the data set, of which the two largest are the following: in the spring of 2020 during a period of 1.5 months the scanner underwent maintenance and was not recording while on 19th December 2021 the scanner stopped recording data unnoticed for a period of 18 days due to the holiday season.

The JarKus data set is collected yearly based on different measurement techniques. Every 250 m along the Dutch coast a transect is measured to determine the current beach and dune profile as well as the current position of the coastline with respect to a reference coastline. In our observation area in Noordwijk two transects (with ID 8200 and 8225) are located (see Figure 4.2A). We focus our comparison on the part on the sandy beach in the aeolian zone. In this part the JarKus data is collected with airborne laser scanning. Findings on these transects, interpretation of data and summary of most relevant parameters are summarised in reports by the Dutch ministry of infrastructure and water management (Rijkswaterstaat, Rijksinstituut voor Kust en Zee, 2022). A visualisation of the trend of the momentary coastline (MKL) in 2022 is shown in Figure 4.3.

Additionally data from weather stations of the Royal Netherlands Meteorological Institute Koninklijk Nederlands Meteorologisch Instituut, 2022 in IJmuiden and Hoek van Holland is used to compare our findings to meteorological properties, such as wind speed and direction and precipitation.

Figure 4.3: Location of JarKus transects around Noordwijk in the Netherlands with current trend of MKL in 2022 as a bar plot at each transect location (Rijkswaterstaat, Rijksinstituut voor Kust en Zee, 2022).

4.3. GENERATION OF INVENTORY OF TRENDS

In order to simplify the 3D point clouds, we rasterize each one into a grid. Then we generate the so called *inventory of trends* out of time series at the grid locations on the beach. The resulting database of trends is then used for further analysis of specific regions and time periods as well as to come up with overall statistics of the entire field site.

The method is explained in several steps as visualised in Figure 4.4. In a first step a data set of corrected point clouds is generated from the raw data, see Section 4.3.1. Then, digital elevation models (DEMs) are used to derive a set of elevation time series located at regular grid points, as discussed in Section 4.3.2. With the help of these time series and multiple hypothesis testing (MHT) a trend inventory is built, Section 4.3.3, which allows then to derive the results for interpretation of natural and anthropogenic processes, compare Section 4.3.4.

4.3.1. CORRECTED POINT CLOUDS AND FILTERING

The laser scanner data is in a local coordinate system with the location of the laser scanner as its origin. The height of the scanner was determined when mounted with a GPS receiver. This constant elevation is added to the z-coordinate, to process actual height above sea level instead of negative elevation with respect to the scanner's location.

In a next step the inclination data from the laser scanner is averaged for each scan and the mean pitch and roll inclination are extracted. These values are then used to calculate rotation matrix R, given by:

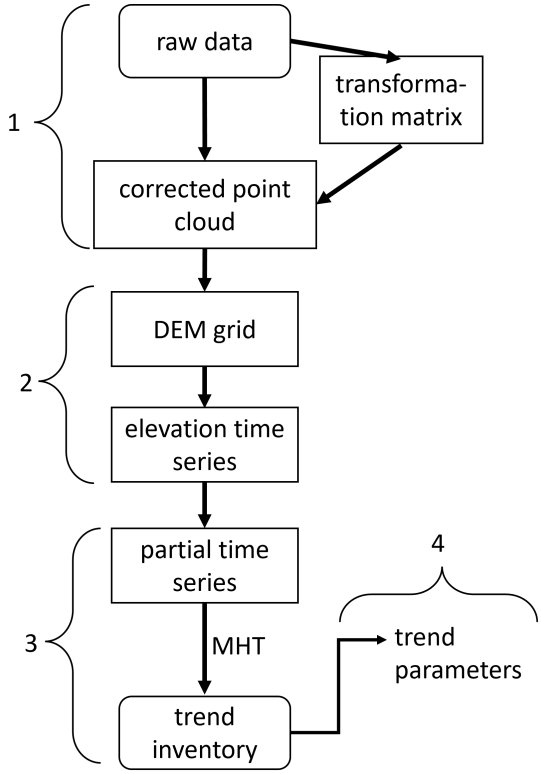

Figure 4.4: Flow chart of the methodology. The numbers refer to the four data processing steps, 1: generation of corrected point cloud, 2: reformatting as elevation time series, 3: data reduction to inventory of trends and 4: Reporting of statistical trend parameters.

$$R = \begin{pmatrix} \cos(\varphi) & 0 & \sin(\varphi) \\ \sin(\varphi)\cdot\sin(\theta) & \cos(\theta) & -\cos(\varphi)\cdot\sin(\theta) \\ -\sin(\varphi)\cdot\cos(\theta) & \sin(\theta) & \cos(\varphi)\cdot\cos(\theta) \end{pmatrix} \tag{4.1}$$

with φ the pitch angle and θ the roll angle. Rotation matrix R is calculated for each point cloud separately and applied, if the standard deviation of the pitch and roll values during the entire scan is below a threshold. This is to ensure that data that was e.g. obtained during a heavy storm is not corrected with a rotation based on erroneous inclination values. If the threshold is not met, a mean pitch and roll value based on the other scans in that month is used for correction. A constant general tilt of the point cloud is removed in this way, as well as the main part of deviations of the laser scanner's position due to temperature changes (heat expansion of supporting materials) or strong winds. We do not use any ground control points for correction of individual point clouds. For validation, some ground control points were measured at the start of data acquisition.

Subsequently the data set is filtered for low quality or erroneous point clouds. The mean elevation as well as the sum of residuals deviating from a fitted plane through a reference surface are estimated for each point cloud. If the mean elevation deviates more than 10 cm or the summed residuals are larger than 10 cm the point cloud is not considered for our study. This threshold was determined during a previous study, (Kuschnerus, Lindenbergh, Vos, & Hanssen, 2024), considering the error sources and influences on data quality. This is to exclude point clouds with large variations/noisy data caused by instrument failure or storms. Point clouds recorded during low visibility weather conditions, which do not contain data on the sandy beach, are excluded from further analysis as well. This leaves a data set of 19 386 good quality point clouds acquired between 11 July 2019 at 14:00 and 21 June 2022 12:00 (midday).

4.3.2. DIGITAL ELEVATION MODELS AND ELEVATION TIME SERIES

The corrected point clouds are then cropped to the relevant part to contain only the sandy beach. For further processing a grid of 1m × 1m is formed and all elevations in each grid cell are averaged to obtain an elevation at the centre of the grid cell, compare Figure 4.5. The 1m × 1m-grid size is a common standard for analysis of coastal spatial data and models. Additionally, it was chosen in order to be as small as possible, while still containing enough points to derive statistically significant mean elevation on most locations in the study area. A sensitivity analysis on point clouds from our data set in comparison with higher resolution data at the same location was presented by Di Biase *et al.*, 2022. The square root of the summed residuals are calculated for each grid cell:

$$\sigma_g = \sqrt{\frac{\underline{\mathbf{e}}^T \underline{\mathbf{e}}}{n_p - 1}}, \tag{4.2}$$

for the spatial residual vector $\underline{\mathbf{e}}$ and number of points n_p. The summed residuals are then saved for each grid cell, as well as the number of points per grid cell. Since the elevations in the DEMs are as a consequence available at regular grid points, all files that meet quality criteria are then collected to generate a time series of 19 386 epochs for each grid cell location. For each grid cell location we have the elevation time series with the mean elevation per time stamp and the accompanying uncertainty time series with the summed residuals per time stamp. Each time series may contain gaps for several reasons:

- a point cloud was not available because it was not recorded or of bad quality,
- an object or pile of sand is temporarily blocking the view on the grid cell, or,
- the tides cause the grid cell to be temporarily under water.

That last case is used to determine events, where high water reaches the study area. For every day we determined the point cloud with the narrowest beach width and save that moment of time and water line at that time as our estimation for the high water line on that day and the time of high tide. This method was described in more detail in a previous study to determine the development of the beach width (Kuschnerus, Lindenbergh, & Vries, 2022).

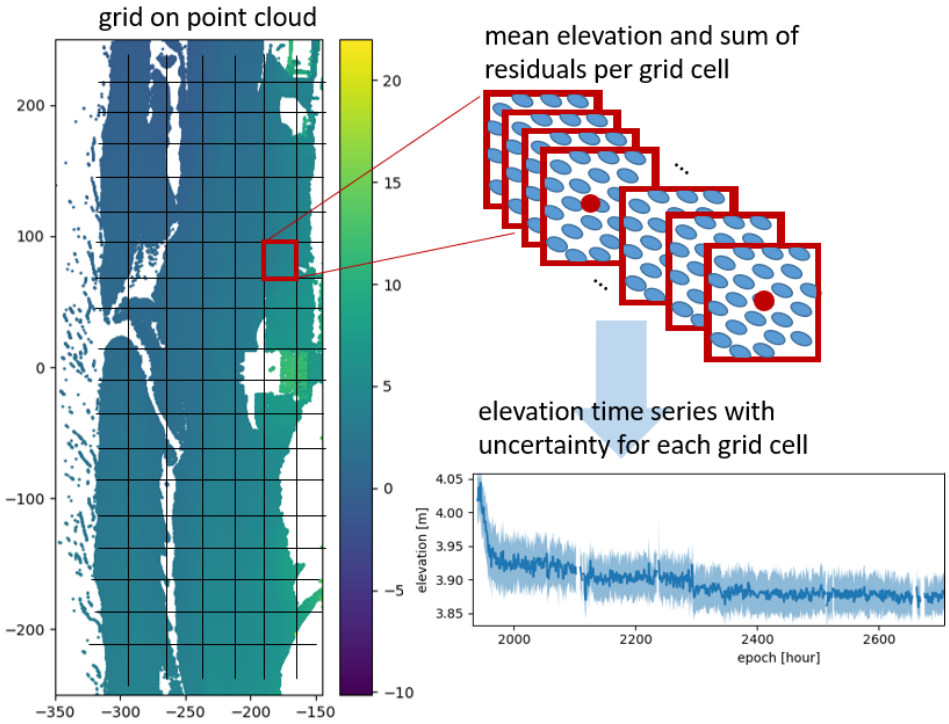

Figure 4.5: Schematic of time series generation: each point cloud is divided into grid cells. For each grid cell all recorded points are summarised by mean elevation and sum of residuals. These values for all available point clouds then provide a time series of about 19 000 elevation values with associated sum of residuals for each grid cell location.

4.3.3. 4D INVENTORY OF TRENDS OF SURFACE ELEVATIONS

Each time series is divided into partial time series. A break point is recorded when there is a gap in the time series of more than three hours or if a jump is recorded in the time series. To detect the jumps or change point the so called Pruned Exact Linear Time (PELT) algorithm is used, (Killick *et al.*, 2012). The PELT algorithm is using minus the maximum log-likelihood as a cost function and adds a point as change point, if it reduces the segmentation costs by more than a penalty term. We used the implementation of the algorithm in the *ruptures* Python package, (Truong *et al.*, 2020). This will also split partial time series with non-monotonic trends at the points where the trend changes.

After the time series is divided into partial time series we use multiple hypothesis testing (MHT) to build the 4D inventory of trends. Goal is to classify each partial time series as (i) stable, (ii) linearly increasing or decreasing, or, (iii) without matching model.

As described in detail in Kuschnerus, Lindenbergh, Vos, and Hanssen, 2024 and similar to Lindenbergh and Hanssen, 2003, we use MHT based on the method presented

by Chang and Hanssen, 2016. Chang and Hanssen, 2016 use MHT to determine an appropriate deformation model for time series of kinematic parameters from InSAR data. Here, we are interested in estimating linear trends for each of the partial time series. Therefore, we choose a linear model as alternative hypothesis to the null-hypothesis. The null-hypothesis, H_0, states that there is no statistically significant change in the respective time series segment.

We first test for the null-hypothesis (i.e. no statistically significant change). If the null-hypothesis is accepted, we save the start and end of the partial time series, slope value 0 and mean elevation over the entire duration. The null-hypothesis is formulated as follows:

$$H_0 : A_0 = \begin{pmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{pmatrix}, \quad x = \underline{\hat{h}} \tag{4.3}$$

$$T_0 \sim \chi^2(1,0), \tag{4.4}$$

For model matrix A_0 and test value T_0, which is centrally χ^2-distributed with 1 degree of freedom, since we have one free parameter that is estimated. The free parameter is here the mean elevation value, $\underline{\hat{h}}$, of the time series of elevation values $\underline{h}$.

As alternative hypothesis, H_1, we consider a linear trend:

$$H_1 : A_1 = \begin{pmatrix} 1 & t_1 \\ \vdots & \vdots \\ 1 & t_m \end{pmatrix}, x = \begin{pmatrix} h_0 \\ v \end{pmatrix} \tag{4.5}$$

$$T_1 \sim \chi^2(1,\lambda) \tag{4.6}$$

where h_0 is the intercept and v the rate of change of the linear model and therefore $A_1 \in \mathbb{R}^{m\times 2}$. The test value here follows a χ^2-distribution with 1 degree of freedom and non-centrality parameter $\lambda \neq 0$, since we have a 2-dimensional model matrix and one additional free parameter compared to the null-hypothesis. The covariance matrix Q_h is defined using the RMSE per grid cell $\underline{\sigma_g^2}$, which were saved for every grid cell together with the elevation time series. We assume that there is no correlation between subsequent measurements and therefore obtain the diagonal matrix

$$Q_h = \text{diag}(\sigma_{g1}^2, \ldots, \sigma_{gm}^2) = \underline{\sigma_g^2} \cdot I_m, \tag{4.7}$$

with

$$\underline{\sigma_g^2} = (\sigma_{g1}^2, \ldots, \sigma_{gm}^2) \tag{4.8}$$

the vector of epoch dependent grid-wise variances σ_{gi}^2, for $i = 1, \ldots, m$.

Following Chang and Hanssen, 2016, the test value T_1 for the alternative hypotheses, is then calculated as

$$T_1 = \sum_{i=1}^{m} \frac{(h_i - \hat{h})^2}{\sigma_i^2}, \tag{4.9}$$

for modelled time series $\underline{\hat{h}} = A_1 \cdot \hat{x}$. The test value is then compared to the critical value k_α, determined by significance level α and non-centrality parameter λ. The chosen parameters in our case are significance level $\alpha = 0.95$ (determining a false alarm rate of 5%) and the power of the test $\gamma = 0.8$ (80% probability of missed detection). Since we are using the χ^2-distribution, the non-centrality parameter λ needs to be chosen in order to represent the significance value α and the power of the test γ. More details on the determination and influence of λ can be found in Kuschnerus, Lindenbergh, Vos, and Hanssen, 2024. Here we use $\lambda = 7.85$ to represent a statistical test set up with our chosen parameters, i.e. 5% probability of detecting a trend, even though there is none, and 80% probability of detecting a trend, when one is present.

If a statistically significant trend is detected, we save the value of the trend in m/day together with the start and stop time of the partial time series and the initial elevation (estimated as axis-intercept). In this way we derive a table of start and stop times and slope values for each grid cell location.

4.3.4. PARAMETER EXTRACTION

The trend inventory delivers a database of statistically significant trends for each time step and each location in our field site. For specific questions, for example the impact of one storm, we can now zoom in on a time period and collect all available partial time series, the matching trends, and the overall change in volume. Also, specific areas can be highlighted, observed over the entire observational period as well as for shorter time spans. To get a quick look on what could have happened for different zoom-in situations (in space and/or time), we calculate the total number of partial time series, the number of partial time series with significant trends, the average duration of the partial time series with significant trends and the average rate of change from the trend inventory. Additionally we report on net volume and area change over the time period, as well as the total sum of absolute volume change. Together, these parameters provide indications on the net results of erosion and deposition processes as well as a measure of how dynamic an area has been.

4.4. RESULTS

As results of our analysis we present statistics on the inventory of trends of the entire area (section 4.4.1) and a comparison with transects from the JarKus data set (section 4.4.2). Then we show how some commonly occurring physical processes appear in the trend inventory (section 4.4.3). Finally we compare the trend inventory for an area with frequent bulldozer works with an area that is dominated by natural processes (section 4.4.4).

4.4.1. INVENTORY OF TRENDS

number of grid cells	41 000
number of partial time series	12.8 million
average duration of partial time series	45 hours
number of significant change rates	395 825
average duration of ts with significant change rates	158 hours
max duration of ts with significant change rates	1209 hours
average change rate	0.0027 m/day
number of stable ts	3.2 million

Table 4.3: Summary of statistics of trend inventory. Time series is denotes as 'ts'.

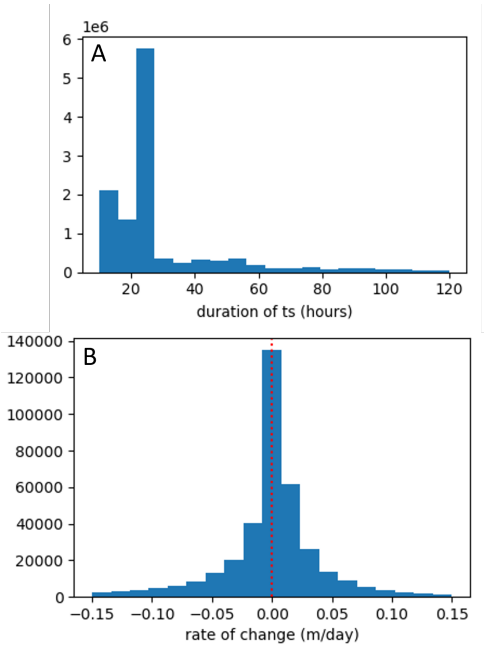

Figure 4.6: Histograms of duration of partial time series (A) and of rates of change (B).

Following the above described procedure, we calculated the inventory of trends for the entire data set covering three years of hourly observations. The statistics on the entire area are listed in Table 4.3 together with the histogram of the rates of change in Figure 4.6. We can see in the table that over the entire area of 41 000 grid cells we have 12.8 million partial time series. For a large part of those time series (about 8 million) we cannot find a statistically significant linear model. The reason for that is either that the shape of the partial time series does not fit a linear model or that the partial time series is too short or has too many gaps in order to be considered. About 3.2 million partial time series appear to represent a period of time where there is no statistically significant rate of change in elevation recorded. We also observe that the average duration of partial time series is 45 hours, but the average duration of partial time series with a statistically significant trend is more than three times longer (158 hours), with a maximum duration of 1209 hours (about 50 days). The minimum duration for all partial time series is 10 hours and was fixed as a setting in the processing.

In Figure 4.7 we show a time line of all major storm events and data gaps for three example time series of raw elevation data as well as fitted linear models for the respective partial time series. The elevation at the time of the JarKus data acquisition at the corresponding location is marked with dots in the same colour as the respective time series. Occurrences of measured wind speeds of 18 m/s and higher are indicated with a yellow

bar. The most prominent gaps in the time series are due to failure in data acquisition or bad quality data. Major storms during extreme wind speeds (as for example storms Ciara and Dennis in February 2020 and Eunice in February 2022) can be observed as well at times where there are abrupt changes or gaps. Another possible reason for sudden changes or jumps is a beach scarp that form on the beach as shown in Figure 4.2B. We observe lots of dynamics (up to 1 m in elevation difference) in between two JarKus height estimates, which are missed by interpolating data from yearly measurement campaigns for JarKus.

4.4.2. COMPARISON OF TREND INVENTORY WITH JARKUS TRANSECTS

We use the JarKus data set as a reference since it is widely used to monitor coastal dynamics. We compare the JarKus transects in our study area with the trend inventory at the same locations and over the same three-year observation period.

After conversion of the coordinate systems we can place a JarKus transect on a point cloud collected on the same date and in this way verify, that the profiles and locations largely match. The transects are orthogonal to the *basic coastline* (BKL) and cover several cross sections of our grid structure. That is why we use an area of 10 m width to cover each of the two JarKus transects (see Figure 4.2). Figure 4.8 shows the JarKus transects at the location with ID 8225 together with a transect of a point cloud from the PLS data set acquired on the same day. We use this to match the JarKus transects with our data and validate the NAP elevation used for the calculation of the trend inventory. We are able to show the variability in the same locations as the transects in between measurement campaigns (as shown in Figure 4.7).

However, in Figure 4.8 it becomes clear that the JarKus transects cover a much larger area across shore, with large parts under water, while our study area only represents a small part of the transects. Any conclusions that we can draw out of our data set are therefore difficult to compare and match with the results on long-term coastal development from JarKus data. JarKus data is used for a variety of applications including the determination of the so-called *MKL-trend*, the trend of the momentary coastline (MKL), describing the current state of the coast compared to the previous years. A similar analysis with our data set would be based on a much smaller area, covering only the aeolian zone, as well as a relatively short time-span of three years. Therefore it proofs to be challenging to draw direct conclusions out of the statistics of the inventory of trends that are comparable to the MKL-trends based on JarKus measurements.

4.4.3. CATEGORISATION OF DIFFERENT TOPOGRAPHIC CHANGE PROCESSES

To show in more detail the potential of the information captured by our method, we discuss how three common process categories, a major storm, aeolian sand transport and bulldozer work, show up in our trend inventory. This demonstrates how our trend inventory can be used to analyse dynamics in between the yearly JarKus measurements with respect to different processes and at different time scales.

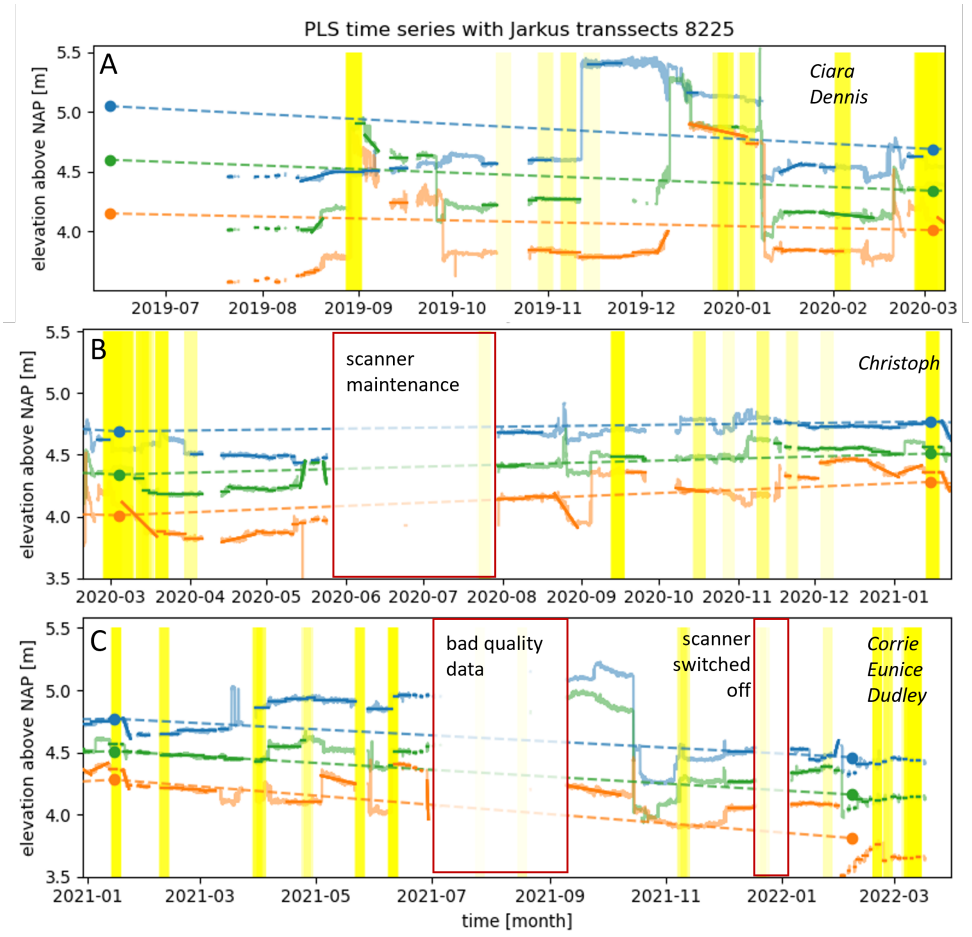

Figure 4.7: Time line of data collection split up in the periods between JarKus campaigns (A: 2019/2020, B: 2020/2021, C: 2021/2022) for three example grid cells along the JarKus transect with ID 8225. JarKus data collection is indicated by a dot at the acquisition time and measured elevation of the transect. The raw elevation data is shown for three grid cells (blue, orange and green time series) along JarKus transect 8225. The fitted trends are shown in the same colour on top of the raw data. Hours with an average wind speed ≥ 18 m/s, are indicated by a yellow bar. The names of some prominent major storms are provided in italics and reasons for large gaps in the data are given in/next to a red box.

MAJOR STORMS

In Figure 4.9 we show the effects of two major storms, Corrie and Eunice/Dudley in February 2022 exemplary at 10 grid cells. The example grid cells are located along a line perpendicular to the coast at $y = -55$ and $-215 \leq x \leq -197$ in steps of 2 m. The bottom

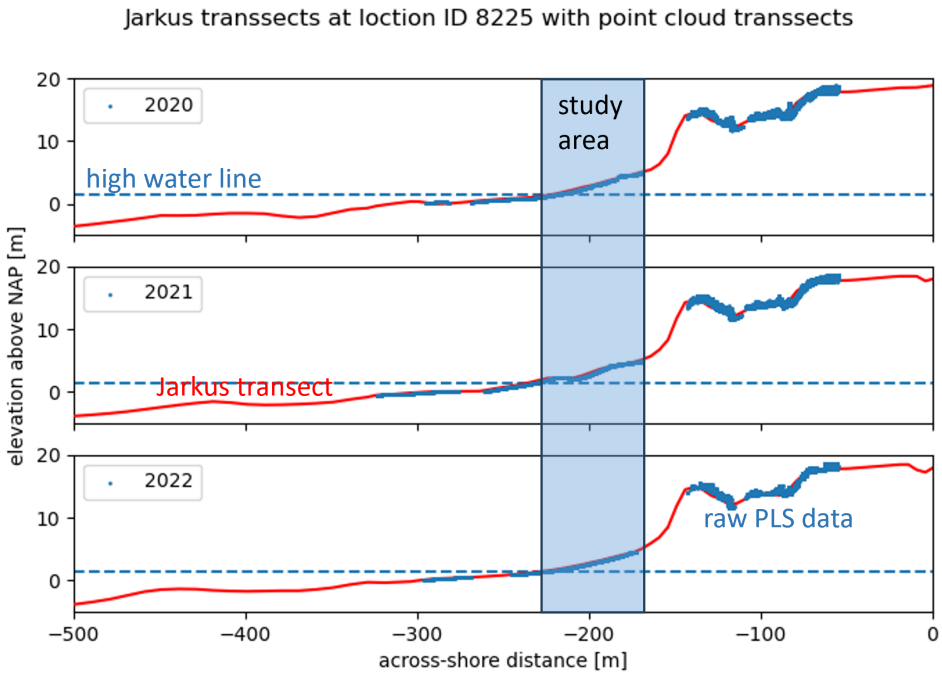

Figure 4.8: Jarkus transects for 2020, 2021 and 2022 with the transect from a point cloud from raw PLS data acquired on the same day. The high water line and the borders of our study area are shown as well.

panel shows wind speed at the same time with a threshold of 18 m/s as indication of a 'major storm'.

Storm Corrie would not register as a major storm according to our wind data collected at IJmuiden weather station. However, it occurred together with high tide resulting in an exceptional high water level which caused large parts of the beach to be temporarily flooded. In Figure 4.9 A, the blue bars indicate where the high water reached up to the location of the grid cells of the respective time series/fitted trends. Since all our study site is in the aeolian zone, flooding during tides happens only mostly during storm conditions.

Eunice and Dudley were overlapping in time and caused very strong wind gusts but we did not register high water reaching our study area. Both storms caused some loss of data, either because of tides covering parts of the beach or through decreased data quality during strong winds. With the trends at different locations on the beach, we can clearly follow the erosion patterns that are caused by both storms. Possible reasons for erosion are either wave run up reaching the study area in between scans, high water covering the area during gaps in the data or aeolian sand transport. We record high water reaching the study area, when the smallest point cloud in one day shows the high water line within the study area. It is possible that these events happen without being recorded

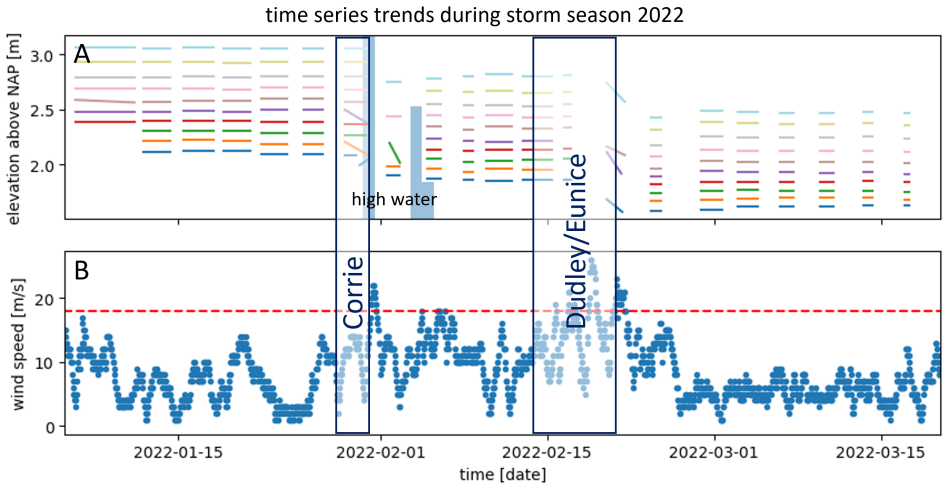

Figure 4.9: A: Time series trends at 10 different grid cells indicated by 10 different colours in storm season 2021/2022 covering the major storms Corrie, Dudley and Eunice. B: Wind speed in the same time period. The dotted red line at 18 m/s wind speed provides the threshold for a 'major storm'. Occasions, where high water reaching up to the grid cells of the shown time series was recorded in our data set, are indicated with a blue bar.

or that wave run up reaches the area while the high water line is just below the border of the study area.

Overall we estimate for the entire observation area a total positive volume change of 1667 m^3 for storm Corrie (with a duration of three days) and a total negative volume change of -3459 m^3 for Eunice/Dudley during six days. Our full observation area, compare Figure 4.2, is roughly 41 000 m^2, thus resulting in an average gain of 4 cm per square meter during storm Corrie and an average loss of 8.4 cm per square meter during storms Eunice and Dudley. Most of the accretion is happening during data gaps, but we suspect that high water reaching the aeolian zone as well as wave run up are the main causes for accretion. We remark here that the impact of both storms was very different and even contrasting. The elevated water levels caused by Corrie are a likely explanation, despite the corresponding relatively lower wind speeds.

AEOLIAN SAND TRANSPORT

Additionally to large and sudden changes, the trend inventory is suitable for the observation of slow elevation changes developing during several hours or even days. For an example of aeolian sand transport we show time series trends at nine locations along a profile perpendicular to the coastline ($y = 150$ and $-206 \leq x \leq -198$) in steps of 1 m together with recorded wind speeds at that time (Figure 4.10). The slow but continuously observed erosion over several days in combination with wind speeds higher than 8 m/s

lead to the assumption that aeolian sand transport is the most likely cause of the elevation changes in these neighbouring grid cells. According to Rijn, 2022, dry sand with wind speeds of 8 m/s and higher provides the conditions for possible aeolian sand transport on the Dutch coast . We did not report on precipitation here, because precipitation varies locally and data from the weather stations at several 10s of kilometres distance is not very accurate especially in coastal areas.

Figure 4.10A shows how the slow erosion of a few centimetres over the course of 10 to 20 days is captured very well by the inventory of trends. Filtering the entire inventory of trends for partial time series with statistical significant rate of change, v, in the range $0\ m/h < v \leq 0.05\ m/h$, that last at least 6 hours, provides an estimate of all the potential times and locations that aeolian sand transport caused changes in the sediment surface. This amounts to just under 395 000 partial time series with potential aeolian sand transport causing a deposition of sediment of about 1354 m^3 in our study area over three years time. This is equivalent to an average gain in elevation of 3.3 cm per square meter over three years.

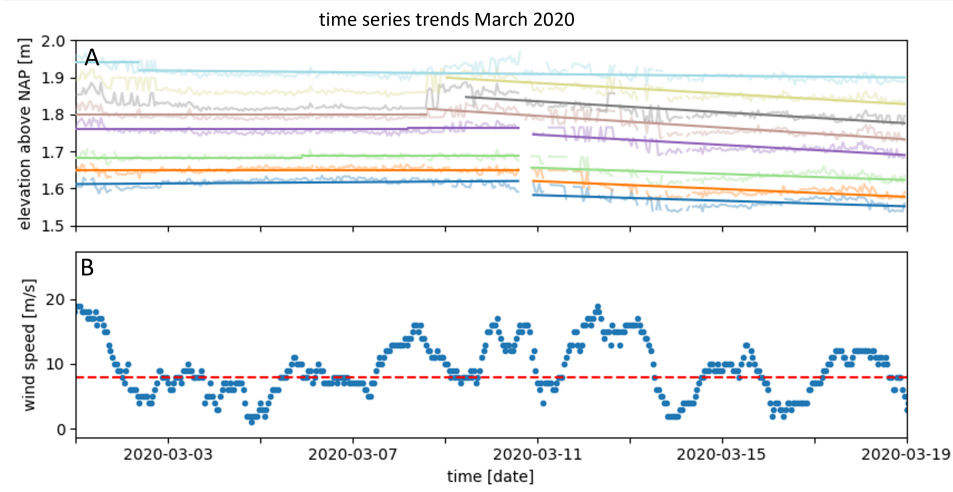

Figure 4.10: A: Time series trends in March 2020 showing most likely occurrence of aeolian sand transport detected through the slow but continuous changes over several days together with wind speeds above 8 m/s (red dotted line, B). The original time series are shown together which each fitted trend.

BULLDOZER WORKS

Our study site is subject to frequent bulldozer works, compare Barbero-García *et al.*, 2023 and Kuschnerus, Lindenbergh, Lodder, *et al.*, 2022. Reasons for bulldozer works are mostly clearance of access paths and preventing and removing storm damage. Generally the beach is maintained frequently to facilitate recreational use. For the interpretation of the data and derived trends it is important to take the potential effects of bulldozer

works into account. A prominent example found in our data set is the set up of a beach tent in test area 1 (see Figure 4.2) next to the beach bar. Before the tent was set up in January 2020, a terrace was formed and the entire month saw frequent bulldozer activities in that area. We use this example to showcase how bulldozer works appears in our trend inventory. In Figure 4.11 we show time series trends for 10 grid locations along the cross shore profile $y = 30$ and $-175 \leq x \leq -156$ in January 2020. In the bottom panel we show the total number of sudden changes per day (of 30 cm or higher) recorded in that area, which are most likely caused by bulldozer works. A list of these potential bulldozer activities was generated for Barbero-García *et al.*, 2023, and is used here as reference indication.

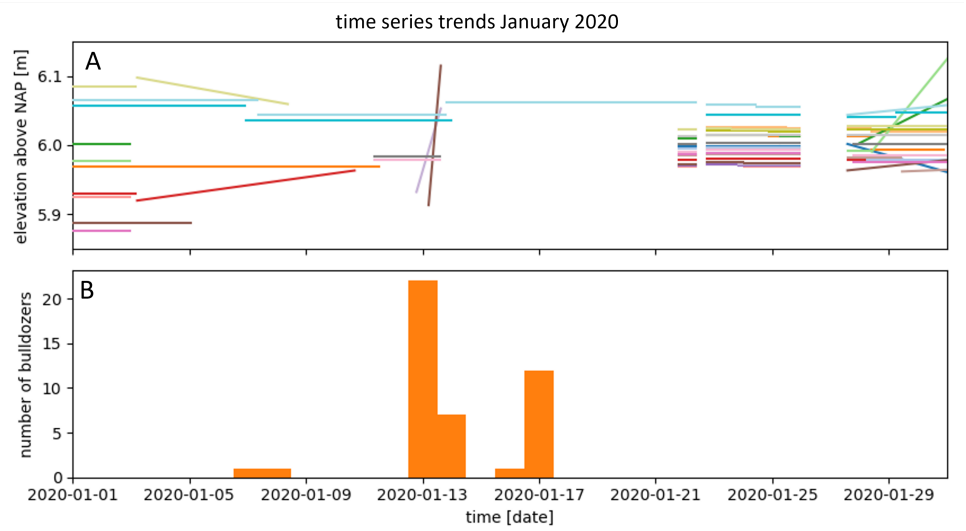

Figure 4.11: A: time series trends for 10 grid locations in January 2020 showing the effects of bulldozer works in the middle of test area 1. B: Number of sudden changes per day. most likely caused by bulldozer works.

We see in Figure 4.11 that the bulldozer works interrupt the ongoing trends significantly. It takes almost two weeks until a regular situation with stable trends appears again. At the same time the bulldozer works lead to loss of data, due to occlusion and frequent interruption of time series. If there are too many gaps in a time series, a trend is not fitted by our method. The tent's set up after the bulldozer works, as well as sand piles can lead to occlusions in neighbouring grid cells.

4.4.4. COMPARISON OF NATURAL VS. ANTHROPOGENIC PROCESSES

Bulldozer works are a common occurrence on the Dutch coast and coasts worldwide but their long-term effects on natural processes are difficult to study.

Here we use the list of instances most likely caused by bulldozer works collected for a previous study (Barbero-García *et al.*, 2023), for a first assessment if a beach is affected

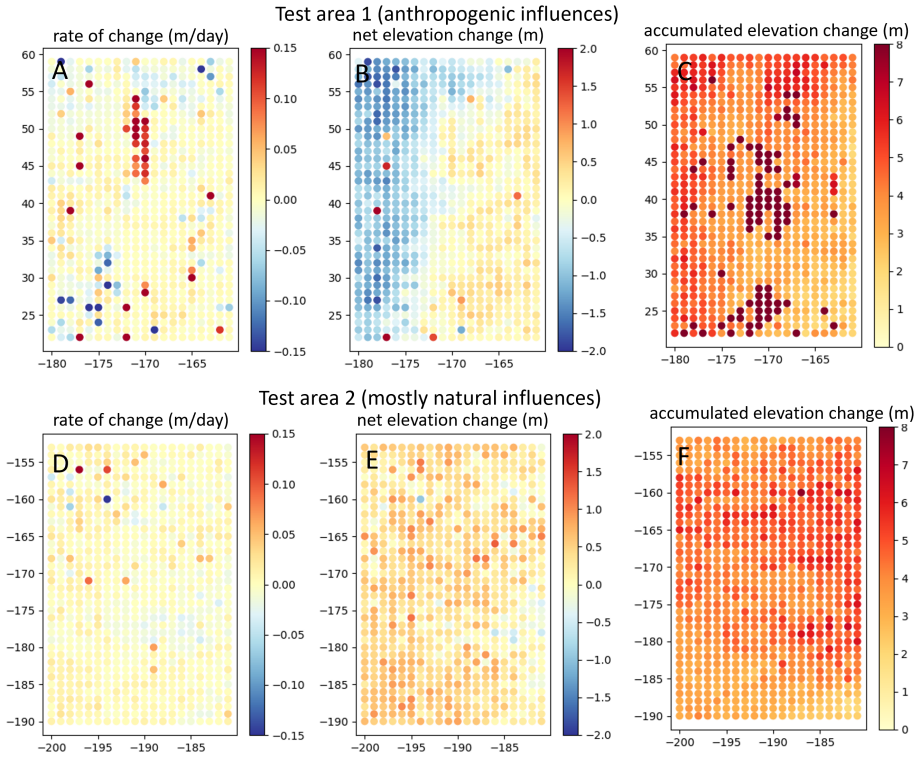

Figure 4.12: Comparison of rates of change, net elevation change and absolute accumulated elevation change over the entire observation period in test area 1 and 2.

by anthropogenic interventions and how many times these interventions occurred. We then chose two dedicated test areas (see Figure 4.2), both of 760 m^2: Test area 1, where we suspect lots of anthropogenic influence (354 recorded incidences of possible bulldozer works) and Test area 2, with few bulldozer works (95 recorded incidences of possible bulldozer works) and therefore mostly natural influences but a similar location on the beach, close to the dunes.

For the comparison of these two areas we show in Figure 4.12 the average rate of change of all fitted trends within the three-year period in that area (A and D) and the net elevation change that we derive out of all the fitted trends in the area (B and E). The accumulated elevation change (C and F) is calculated as the sum of all jumps that appear in the trend inventory and the recorded slow changes both as absolute values.

Test area 1, with lots of anthropogenic processes, shows lots of variability, see Figure 4.12A-C. Over the entire observation period, the net volume change in that area was -257.5 m^3 with an average rate of change of 0.0071 m/day. Figure 4.12B shows a clear

pattern in both the yearly and entire period accumulated change from the dunes towards the sea, with erosion closer to the sea (west) and no change or some deposition closer to the dunes. The average rate of change however, Figure 4.12A, reveals different patterns.

For comparison, Test area 2 is dominated by natural processes. In Figure 4.12D-F we do not observe the patterns observed for Test area 1: Over the entire observation period the net volume change in Test area 2 was 272 m^3 with an average rate of change of 0.002 m/day. The overall accumulated net change as well as the average rate of change over the entire three-year period are visualised in Figures 4.12E and F. This leads to the hypothesis, that more frequent anthropogenic human activity occurs in areas where we also observe more erosion, compared to areas with less human influences.

To relate the amount of human activity (i.e. bulldozer works) in an area to the variability and erosion/deposition patterns in that area for our entire observation period, we compare the number of potential bulldozer works over the entire three-year period to the variability, quantified as net change: with a moving window of 20 m length across-shore and 40 m length along-shore we collect for each window the number of bulldozer work instances and the net change over the three-year period.

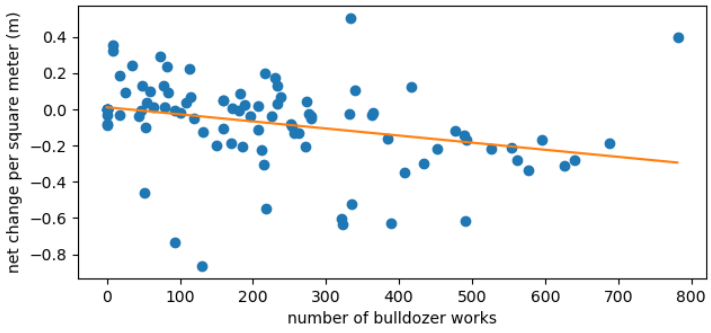

Figure 4.13: Comparison of number of bulldozer works with net elevation change per square meter in the same area. The orange line shows a fitted trend.

The result of this analysis is shown in Figure 4.13, indicating a (small) negative correlation between net change and number of bulldozer works. The correlation coefficient of 0.3 supports this result. We therefore observe, that in our study site, most bulldozer works are happening in areas that were eroding during the observation period.

4.5. DISCUSSION

We conclude that with the large and complex PLS data set, it is possible to observe interacting processes with unprecedented detail and complexity. The comparison with JarKus transects provided the opportunity to validate the elevation in the PLS data set. However, it also showed that a simple question, such as 'Is the coastline at a specific location retreating or advancing?' is not easily answered using our more complex data. For this question the JarKus data set is more suited to provide easily interpreted quantified

answers. Generally speaking, we find a trade-off between the opportunity to research increasing complexity of a natural system vs. the need for clear and easy to communicate answers on major challenges of coastal engineering and monitoring. With the data from the trend inventory, we can provide general indicators of variability and also if there is generally a dominating trend of erosion or deposition. However, a direct comparison with the indicator of for example the MKL-trend is difficult.

The unique nature of the analysed data set lends itself to look for systematic analysis of coastal dynamics, which are mostly ruled by natural processes. The question if a 'heavy storm' causes more surface changes depending on the strength of the wind speed, was considered in this context. However, a consistent and systematic definition of a storm, based on one or two observed parameters, such as wind speed and wind direction proved to be difficult. As can be seen in our analysis (Section 4.4.3 and Figure 4.9) two major storms present themselves differently in terms of wind speed as well as with their impact on the beach. We decided not to present an attempt at correlating for example sand volume change and wind speed in order to avoid a simplified presentation of a complex system of interacting natural forces and their effects. To improve this in similar future experiments, a weather station next to the laser scanner as well as wind meters in several locations on the beach would be recommended.

The PLS data set together with the inventory of trends provides many opportunities for more research. We showed in section 4.4.3 how specific physical processes that cause elevation change at the sediment surface appear in the inventory of trends. Especially storm damage but also the more gradual and difficult to monitor aeolian sand transport appear with a clear pattern. Considering wind speed together with the trend inventory both can potentially be detected in a data set of a large long term study. There is also the opportunity to take a closer look at different regions and/or time scales. For example the analysis of seasonal changes in erosion/accretion patterns on the sediment surface could be interesting. However, accretion seems to happen mostly during storm conditions in our study area. Therefore the trend inventory appears not to be the ideal method for the detection and identification of accreting events. Adding additional means of observing the study area, for example through video cameras, could help to confirm the suspected causes of erosion/accretion during times of data loss and improve this point.

The acquisition of more PLS data sets could lead to the opportunity for the generation of training data for a possible supervised learning algorithm to detect these processes. We still consider our three-year data set as insufficient for training an algorithm on the occurrence of a specific process, such as aeolian sand transport or a storm. But, we feel we have successfully demonstrated the possibility of identification of such processes in a PLS data set.

For the evaluation of the impact of anthropogenic activities, data from other locations and possibly specific long term experiments could be useful. The slight correlation between erosion and amount of bulldozer works that we show in section 4.4.4, could be an indication, that frequent human interventions disturb the naturally occurring deposition of sand. Another interpretation could be that larger erosion requires more corrective bulldozer interventions. In any case it is important to consider the context, that all bulldozer works must be below 20 m^3 per transect of 1 m width, according to the Dutch regulations. This threshold is chosen in order to ensure that bulldozer works are

in the same order of magnitude as natural processes and do not disturb them. However, the correlation that we show, could be an indicator that either more bulldozer works are happening than the regulation allows, or the threshold was based on a false assumption and there is indeed an impact of the frequent anthropogenic activities on natural processes. Both of these hypothesis need further research.

4.6. CONCLUSIONS

We analysed a three-year data set from PLS with the help of hypothesis testing, to find statistically significant monotonic trends in elevation and reduce the complex data set to an inventory of trends. We furthermore showcased applications of the inventory of trends for the analysis of change patterns at the sediment surface in the aeolian zone of a Dutch sandy beach by answering four research questions:

- What elevation trends at the sediment surface can be derived from the PLS data set?

  We found about 380 000 thousand partial time series with statistically significant rates of change of 0.0027 m/day on average with an average duration of 158 hours. We therefore show dynamics on the sediment surface that are difficult to study and to reveal with other techniques. Especially slowly developing elevation changes over longer periods of times (several days/weeks) and at cm level can be captured with our method.

- How do high resolution PLS data compare to yearly JarKus-data?

  The elevation from the PLS data set is validated at two locations of the JarKus transects at the day of acquisition of the JarKus data. Therefore our data set and method allows to show dynamics in between the measurement campaigns at these specific locations and in a small part of the JarKus transects. The comparison of long-term (up to 10s of years) trends concerning the coast line proofs to be difficult. Our more complex and exhaustive data set, on a smaller area does not allow to draw simple conclusions comparable to coastline trends from JarKus data, but shows dynamics that are missed by the yearly JarKus measurements.

- How can small and slow changes at the sediment surface be identified and quantified with PLS?

  We show that large elevation changes, for example caused by storms as well as slow, small-scale changes in elevation over longer periods of time can be identified and analysed with the inventory of trends. Especially small-scale, long term elevation changes, which are associated with aeolian sand transport are well captured within the inventory of trends. Anthropogenic activities such as bulldozer works can be found as well, but are easier detected with other methods, for example by considering sudden changes as presented by Barbero-García *et al.*, 2023.

- How can the effect of anthropogenic activities on the dynamics at the sediment surface be analysed using PLS data?

By comparing the number of bulldozer sightings in an area with the net elevation change in the same area and time frame we find a connection between erosion and bulldozer works with a correlation coefficient of 0.3. This leads to the hypothesis that bulldozer works negatively affect natural processes on the beach and might lead to more erosion. Further research is needed to confirm this hypothesis.

We presented a new method to gain information on change processes at the sediment surface on a sandy beach from a PLS data set. To extend this research, we suggest comparison with similar data sets at other locations, as well as collection of more environmental variables (for example in situ wind speed and direction) in future similar experimental set-ups. The trend inventory is a tool that could potentially be used for many other applications and case studies, such as analysis of seasonal patterns in erosion/deposition, study of intertidal bar dynamics or dune growth.

5

CONCLUSIONS AND RECOMMENDATIONS

This dissertation deals with the development of methods for spatio-temporal data mining of a large 4D data set from permanent laser scanning (PLS), to answer the following research question:

How can geomorphologic processes be identified and quantified in a 4D spatio-temporal data set from permanent laser scanning?

5.1. CONCLUSIONS

To answer the main research question the quality of two PLS data sets was analysed and the main influences on systematic and random errors on height estimates using PLS were quantified. Further, two suitable methods to detect and identify geomorphologic changes in a 4D spatio-temporal data set from PLS were provided. These methods allow to analyse specific processes on the sediment surface and to provide quantified results. Clustering time series allows to group the entire area into different zones according to deformation behaviour at the sediment surface. By clustering time series, processes with a specific change pattern at a particular time stamp are grouped together and allow to identify for example bulldozer works, without defining the expected change pattern in advance. The second method uses multiple hypothesis testing to detect changes following a given model time series, for example linear changes or sudden jumps, with predefined detection power and statistical significance. With the help of MHT an inventory of trends was generated and a means to process a large PLS data set covering several years was developed. The inventory of trends is particularly suitable to find slow gradual processes, which are mostly caused by aeolian sand transport. Additionally, bulldozer works and storm effects can be detected using this method. Comparing the areas influenced by bulldozer changes and the erosion occurring in these areas, correlation between number of occurrences of bulldozer works and amount of erosion was found.

More specifically, the following conclusions were found answering the different subquestions:

(Q1) *What influences the uncertainty of height estimates from permanent laser scanning on a sandy beach?* (**Chapter 2**)

A 4D spatio-temporal data set from PLS is subject to many different environmental and instrumental conditions, which influence the data quality. In our data set acquired at the Dutch coast in Noordwijk and Kijkduin, instrument failure and technical problems during operations caused the largest data gaps. A technical problem with the laser scanner also caused lower quality data, by adding outliers without physical representation in the real world. The stability and automation level of the instrument set-up, are therefore the most important influences on the continuity and quality of the resulting data set. Environmental conditions such as heavy rains and strong winds, as well as fog have a strong influence by affecting the stability of the system and the visibility of the observed area. Small elevation errors of up to 1 cm per day, as an effect of daily temperature variations, can be observed as well but have less effects on the derived output products (i.e. elevation time series, detected changes, clusters). Considering each location within the observed area, geometric properties and local surface properties, mainly reflectivity and surface roughness are the largest factor of influence, which are however more difficult to quantify and verify independently. Their combined effect are estimated to result in up to 7 cm of random error. The (fine-) registration method used to align all point clouds has a considerable influence as well and is subject to further research. The currently used registration methods are not as exact as desirable and lead to registration errors of up to 1.5 cm. Other commonly used methods, such as iterative closest point (ICP) suffer from a lack for stable reference surfaces and turned out to be too computationally expensive to apply to the entire data sets.

(Q2) *Which change patterns at the coastal sediment surface can be identified with permanent laser scanning?* (**Chapter 2 and 4**)

Theoretically, surface elevation changes on the sandy beach can be detected with a specified probability, if their magnitude is larger than the minimal detectable bias (see chapter 2) and the assumed model matches the respective change pattern. The minimal detectable bias for sudden jumps and linear trends is in the order of centimetres given a discriminatory power of 80%. Additionally, the detection potential varies depending on location and distance to the laser scanner and the length of the considered time series. The underlying processes causing the surface elevation change can be natural or anthropogenic. Many anthropogenic activities are relatively easy to identify, due to sudden occurrence and 'unnatural' characteristics, mainly sudden jumps in elevation. Natural processes with large energy, such as storm surges can be detected and identified as well. Long-term, but small-scale and slowly developing processes are detected and well represented within our inventory of trends (see chapter 4). Comparing the approximated time series from the inventory of trends with weather data (e.g. wind speed) and with daily location of the high water line, allows identification of the most likely underlying process. Aeolian sand transport, as the most common natural process in the aeolian zone, is successfully detected in this way. Geomorphologic changes resulting from tidal forces in the intertidal zone can be identified as well. Here the frequent gaps in the time series and the high variability cause most issues for our methods (clustering, as well as multiple hypothesis testing (MHT)).

(Q3) *How can elevation time series resulting from the same process be grouped together?* (**Chapter 3**)

Using an existing clustering method, time series from a 4D spatio-temporal data set can be grouped together to identify zones with similar dynamic behaviour. The main evolution of the respective zone is following a representative change pattern, which can be analysed to identify the most likely underlying physical process. Several clustering algorithms are suitable for this application. The k-means algorithm with a predefined number of clusters k and DBSCAN turn out to be most suitable for the detection of larger areas behaving in similar ways (k-means) and for unusual changes affecting only a few time series (DBSCAN). Local approximation of time series with the help of hypothesis testing allows to generate an inventory of trends. The inventory of trends provides characteristics to group time series and parts of times series in order to identify dominating processes in specific regions. The start and stop times of each time series piece is estimated as well and provides insight into moments in time where processes are interrupted.

(Q4) *How are anthropogenic activities detected with permanent laser scanning and what is their influence on coastal development and natural dynamics?* (**Chapter 4**)

Identifying all sudden changes allows to detect sand displacements likely caused by human activities. This can be done via a fixed, pre-defined threshold (see Barbero-García *et al.*, 2023). Additionally, with hypothesis testing for sudden changes in the form of a step function model, most of the observed anthropogenic changes are identified. The long term impact of these changes is analysed with the help of the inventory of trends and quantified in form of a mass budget in areas subject to frequent anthropogenic activities. It is concluded that a large part of the effect of a building on coastal development is caused by the accompanying bulldozer activities and not only by the building itself. A slight correlation between number of bulldozer activities and erosion was found, which indicates a relation between the two that goes beyond what is accounted for by Dutch coastal management regulations.

5.2. MAIN CONTRIBUTIONS

1. The unique PLS data set from Noordwijk was processed in its entirety covering 20 875 epochs of hourly acquired point clouds over the course of three years for the first time and a shorter PLS data set from Kijkduin covering six months. The structure as well as large size of the data set make it computationally challenging to process the entire data set. The methods and work-flow developed in this thesis allowed for a restructuring of the PLS data set into gridded time series, which were then processed using multiple hypothesis testing to generate an inventory of trends for easier data access and interpretation (chapter 4).

2. The analysis of the influencing factors on the quality of our data set in general and on height estimations in particular as presented in chapter 2, provides opportunity for future improvements in similar PLS data acquisition set-ups. On one hand

it is concluded that more additional environmental data, like data recorded with an in situ weather station, is needed to provide quantification on the uncertainties of height estimations. On the other hand, the major influencing factors could be identified, and it was shown that for several applications the current data quality as well as error estimations as presented in chapter 2, are sufficient to generate many promising results. Most dominating change patterns could be identified with statistical significance even with sub-optimal registration.

3. Clustering methods were tuned and applied to the use on elevation time series from a PLS data set. The clustered time series, where each cluster is described by one mean time series and the locations of the associated time series provide a promising method for the detection of change patterns and the identification of underlying physical processes (see chapter 3). This approach was taken up by Czerwonka-Schröder, 2023 and applied for determining surface activities on an Alpine slope with possible landslide risk, as well as for distinguishing surface types. The work of Winiwarter *et al.*, 2023 builds on this method to cluster time series from PLS data after Kalman filtering and feature detection to distinguish different zones according to change patterns on a rockfall site.

4. Another way to simplify the information contained in the PLS data set, is the use of multiple hypothesis testing. Multiple hypothesis testing was used to fit different temporal models to partial time series from the PLS data set (see chapter 2). This method allows generating the inventory of trends (chapter 4), estimating partial time series start and stop times and linear model coefficients. This is a summary of all statistically significant rates of change that each time series follows over the entire observation period. Subsequently, these pieces of time series can be grouped together, in the future possibly through clustering, or specific regions of interest can be extracted according to spatial location or timing. In this way the following processes were extracted and identified:

   - storm events
   - bulldozer induced changes
   - tidal variations
   - aeolian sand transport

   Each of these processes is associated with a location and timestamp in the time series data set, where the respective process was dominating. This allows for further analysis of the process itself, its effects and its spatial and temporal extent and improves understanding of geomorphologic coastal processes. Especially small-scale long term processes, which are often not captured by in-situ, incidental observation methods, or by large scale satellite observations, can be observed in this way.

5. Using the described methods, evidence for the importance of anthropogenic influences in coastal areas was found. The analysis of anthropogenic changes led to the conclusion, that not only buildings themselves, but also the associated human interventions on the sandy beach around each building have an impact on

coastal morphology and possibly lead to increased erosion. The current regulations of Dutch coastal management allow for a specific threshold of sand volume to be moved per cross section on the beach. A correlation between erosion and small-scale, but frequent bulldozer activities was found. Thus it is suspected that current regulations do not prevent long term effects on the dynamics of the sediment surface in coastal areas.

5.3. RECOMMENDATIONS FOR FUTURE WORK

The presented system of PLS can be installed and used in any location providing a power supply, internet connection or data saving option on site and an elevated location from which the observation area is overlooked. For a permanent monitoring of the entire Dutch coast this infrastructure is not available and the costs would be very high. However, a similar system can potentially be installed on specific sites all around the world, as exemplary case study for the monitoring and analysis of coastal geomorphological processes. Additionally, the presented methods could be applied to a data set from frequent airborne laser scanning acquired either with planes or uncrewed aerial vehicles (UAV). One example of such a data set data set is the JarKus data set (Rijkswaterstaat - Dutch Ministery of Infrastructure and Water Management, 2022), which consists of yearly point clouds of the entire Dutch coast (see also C. O. IJzendoorn *et al.*, 2021). The presented methods can be used to simplify this data set and help with the analysis of observed processes and changes.

Additionally, the presented methods are suitable for the application to any 4D data set from permanent or (nearly) continuous laser scanning, or other 3D observation techniques, like photogrammetry. Out methods allow for post processing and analysis and research of the respective area, but are not yet adapted to real time warning systems, or prediction. A possible suggested line of future research would be to extend the presented methods with short term predictions and derive real-time warning systems.

Further analysis of the inventory of trends and/or clustered time series will allow improved understanding of coastal processes and their interactions. The unique PLS data sets in combination with the presented methods allows for a thorough analysis of local volume and elevation changes. The underlying processes can be derived or assumed in many cases and therefore physical models for geomorphologic processes can be verified and possibly improved. Long term models for example for dune growth could benefit from additional insights gained this way and the incorporation of small-scale processes as observed in this case study can potentially improve overall coastal modelling.

List of Publications

Journal articles

- **M. Kuschnerus**, S. de Vries, J. A. Á. Antolínez, S. Vos, R. Lindenbergh: Identifying Topographic Changes at the Beach Using Multiple Years of Permanent Laser Scanning, Coastal Engineering, vol. 193, no. 104594, October 2024.
- **M. Kuschnerus**, R. Lindenbergh, S. Vos and R. Hanssen: Statistically Assessing Vertical Change on a Sandy Beach from Permanent Laser Scanning Time Series, ISPRS Open Journal of Photogrammetry and Remote Sensing, ISPRS Open Journal of Photogrammetry and Remote Sensing, vol. 11, no. 100055, January 2024.
- I. Barbero-García, **M. Kuschnerus**, S. Vos, R. Lindenbergh: Automatic Detection of Bulldozer-induced Changes on a Sandy Beach from Video using YOLO Algorithm, International Journal of Applied Earth Observation and Geoinformation, vol. 117, March 2023.
- S. Vos, K. Anders, **M. Kuschnerus**, R. Lindenbergh, B. Höfle, S. Aarninkhof and S. de Vries: A High-resolution 4D Terrestrial Laser Scan Dataset of the Kijkduin Beach-dune System, The Netherlands, Scientific Data, vol. 9, p. 191, April 2022.
- **M. Kuschnerus**, R. Lindenbergh, and S. Vos: Coastal Change Patterns from Time Series Clustering of Permanent Laser Scan Data, Earth Surface Dynamics, vol. 9, pp. 89–103, February 2021.

Publications in conference proceedings

- D. Hulskemper, K. Anders, J.A.Á. Antolínez, **M. Kuschnerus**, B. Höfle, R. Lindenbergh: Characterization of morphological surface activities derived from near-continuous terrestrial LiDAR time series, International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences, XLVIII-2/W2-2022, 53–60, December 2022.
- **M. Kuschnerus**, R. Lindenbergh: Quantifizierung von Baggerarbeiten an der niederländischen Küste mit 4D-Laserdaten, DVW e.V. (Hrsg.): Terrestrisches Laserscanning 2022 (TLS 2022), DVW-Schriftenreihe, Band 104, S. 65–76, Augsburg, December 2022.
- **M. Kuschnerus**, R. Lindenbergh, S. de Vries: Assessing Sandy Beach Width Variations on Intertidal Time Scales Using Permanent Laser Scanning, 5th Joint International Symposium on Deformation Monitoring (JISDM), Valencia, Spain, June 2022.

- **M. Kuschnerus**, R. Lindenbergh, Q. Lodder, E. Brand, S. Vos: Detecting Anthropogenic Volume Changes in Cross Sections of a Sandy Beach with Permanent Laser Scanning, International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences, XLIII-B2-2022, 1055–1061, June 2022.
- **M. Kuschnerus**, D. Schröder, R. Lindenbergh: Environmental Influences on the Stability of a Permanently Installed Laser Scanner, International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences, vol. XLIII-B2-2021, July 2021.
- S. Vos, **M. Kuschnerus**, and R. Lindenbergh: Assessing the Error Budget for Permanent Laser Scanning in Coastal Areas, FIG Working Week 2020, no. 10579, May 2020.
- R. Lindenbergh, S. van der Kleij, **M. Kuschnerus**, S. Vos, and S. de Vries: Clustering Time Series of Repeated Scan Data of Sandy Beaches, International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences, vol. XLII-2/W13, pp. 1039–1046, June 2019.

DATA PUBLICATIONS

- S. Vos, **M. Kuschnerus**, R. Lindenbergh, S. de Vries: 4D spatio-temporal laser scan dataset of the beach-dune system in Noordwijk, NL, 4TU.ResearchData, 2023. doi: 10.4121/1aac46fb-7900-4d4c-a099-d2ce354811d2.v2
- S. Vos, **M. Kuschnerus**: CoastScan: Data of daily scans at low tide Kijkduin January 2017, 4TU.ResearchData, 2020. doi: 10.4121/uuid:409d3634-0f52-49ea-8047-aeb0fefe78af

BIBLIOGRAPHY

Aarninkhof, S., De Schipper, M., Luijendijk, A., Ruessink, G., Bierkens, M., Wijnberg, K., Roelvink, D., Limpens, J., Baptist, M., Riksen, M., Bouma, T., de Vries, S., Reniers, A., Hulscher, S., Wijdeveld, A., van Dongeren, A., van Gelder-Maas, C., Lodder, Q., & van der Spek, A. (2019). ICON.NL: Coastline observatory to examine coastal dynamics in response to natural forcing and human interventions, international conference on coastal sediments. *Coastal Sediments 2019*, 412–419. https://doi.org/10.1142/9789811204487_0037

Abellán, A., Calvet, J., Vilaplana, J. M., & Blanchard, J. (2010). Detection and spatial prediction of rockfalls by means of terrestrial laser scanner monitoring. *Geomorphology*, *119*(3-4), 162–171. https://doi.org/10.1016/j.geomorph.2010.03.016

Ackerley, D., Chandler, J., & Bullard, J. (2016). Monitoring dune response to the east coast storm surge of 2013 using laser scanning and SFM photogrammetry. In *Coastal management: Changing coast, changing climate, changing minds* (pp. 545–554). ICE Publishing. https://doi.org/10.1680/cm.61149.545

Anders, K., Lindenbergh, R. C., Vos, S. E., Mara, H., de Vries, S., & Höfle, B. (2019). High-frequency 3d geomorphic observation using hourly terrestrial laser scanning data of a sandy beach. *ISPRS Annals of Photogrammetry, Remote Sensing and Spatial Information Sciences*, *IV-2-W5*, 317–324. https://doi.org/10.5194/isprs-annals-IV-2-W5-317-2019

Anders, K., Winiwarter, L., Lindenbergh, R., Williams, J. G., Vos, S. E., & Höfle, B. (2020). 4d objects-by-change: Spatiotemporal segmentation of geomorphic surface change from LiDAR time series. *ISPRS Journal of Photogrammetry and Remote Sensing*, *159*, 352–363. https://doi.org/10.1016/j.isprsjprs.2019.11.025

Anders, K., Winiwarter, L., Mara, H., Lindenbergh, R. C., Vos, S. E., & Höfle, B. (2021). Influence of Spatial and Temporal Resolution on Time Series-Based Coastal Surface Change Analysis Using Hourly Terrestrial Laser Scans. *ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, *V-2-2021*, 137–144. https://doi.org/10.5194/isprs-annals-V-2-2021-137-2021

Anders, K., Winiwarter, L., & Höfle, B. (2022). Improving change analysis from near-continuous 3d time series by considering full temporal information. *IEEE Geoscience and Remote Sensing Letters*, *19*, 1–5.

Assent, I. (2012). Clustering high dimensional data. *WIREs Data Mining and Knowledge Discovery*, *2*(4), 340–350. https://doi.org/10.1002/widm.1062

Baarda, W. (1968). A testing procedure for use in geodetic networks. *Publications on Geodesy, New Series*, *2*(5).

Barbero-García, I., Kuschnerus, M., Vos, S., & Lindenbergh, R. (2023). Automatic detection of bulldozer-induced changes on a sandy beach from video using yolo algo-

rithm. *International Journal of Applied Earth Observation and Geoinformation*, *117*, 103185. https://doi.org/10.1016/j.jag.2023.103185

Belgiu, M., & Csillik, O. (2018). Sentinel-2 cropland mapping using pixel-based and object-based time-weighted dynamic time warping analysis. *Remote Sensing of Environment*, *204*, 509–523. https://doi.org/10.1016/j.rse.2017.10.005

Bergen, J. van, Mulder, J., Nijhuis, S., Poppema, D., Wijnberg, K., & Kuschnerus, M. (2021). Urban dunes: Towards BwN design principles for dune formation along urbanized shores. *Research in Urbanism Series*, *7*, 101–128. https://www.rius.ac/index.php/rius/article/view/130

Berndt, D. J., & Clifford, J. (1994). Using dynamic time warping to find patterns in time series. *Proceedings of the 3rd international conference on knowledge discovery and data mining*, 359–370.

Bitenc, M., Lindenbergh, R., Khoshelham, K., & Van Waarden, A. P. (2011). Evaluation of a LIDAR Land-Based Mobile Mapping System for Monitoring Sandy Coasts. *Remote Sensing*, *3*(7), 1472–1491. https://doi.org/10.3390/rs3071472

Borradaile, G. J., & Borradaile, G. (2003). *Statistics of earth science data: Their distribution in time, space, and orientation* (Vol. 351). Springer.

Brand, E., Ramaekers, G., & Lodder, Q. (2022). Dutch experience with sand nourishments for dynamic coastline conservation–an operational overview. *Ocean & Coastal Management*, *217*, 106008.

Brasington, J., Rumsby, B., & McVey, R. (2000). Monitoring and modelling morphological change in a braided gravel-bed river using high resolution gps-based survey. *Earth surface processes and landforms: the journal of the British Geomorphological Research Group*, *25*(9), 973–990. https://doi.org/10.1002/1096-9837(200008)25:9<973::AID-ESP111>3.0.CO;2-Y

Callaghan, D. P., Ranasinghe, R., & Roelvink, D. (2013). Probabilistic estimation of storm erosion using analytical, semi-empirical, and process based storm erosion models. *Coastal Engineering*, *82*, 64–75. https://doi.org/10.1016/j.coastaleng.2013.08.007

Campos, M. B., Litkey, P., Wang, Y., Chen, Y., Hyyti, H., Hyyppä, J., & Puttonen, E. (2021). A Long-Term Terrestrial Laser Scanning Measurement Station to Continuously Monitor Structural and Phenological Dynamics of Boreal Forest Canopy. *Frontiers in Plant Science*, *11*. https://doi.org/10.3389/fpls.2020.606752

Çelik, M., Dadaşer-Çelik, F., & Dokuz, A. Ş. (2011). Anomaly detection in temperature data using dbscan algorithm. *2011 International Symposium on Innovations in Intelligent Systems and Applications*, 91–95.

Chang, L., & Hanssen, R. F. (2016). A Probabilistic Approach for InSAR Time-Series Postprocessing. *IEEE Transactions on Geoscience and Remote Sensing*, *54*(1), 421–430. https://doi.org/10.1109/TGRS.2015.2459037

Coppi, R., D'Urso, P., & Giordani, P. (2010). A fuzzy clustering model for multivariate spatial time series. *Journal of Classification*, *27*(1), 54–88. https://doi.org/10.1007/s00357-010-9043-y

Czerwonka-Schröder, D. (2023). Konzeption einer qualitätsgesicherten implementierung eines echtzeitassistenzsystems basierend auf einem terrestrischen

long range laserscanner [dissertation, Technische Universität Bergakademie Freiberg]. https://nbn-resolving.org/urn:nbn:de:bsz:105-qucosa2-857419

Davidson, M., Van Koningsveld, M., Kruif, A. de, Rawson, J., Holman, R., Lamberti, A., Medina, R., Kroon, A., & Aarninkhof, S. (2007). The CoastView project: Developing video-derived Coastal State Indicators in support of coastal zone management. *Coastal Engineering*, *54*(6), 463–475. https://doi.org/10.1016/j.coastaleng.2007.01.007

de Vries, S., Southgate, H., Kanning, W., & Ranasinghe, R. (2012). Dune behavior and aeolian transport on decadal timescales. *Coastal Engineering*, *67*, 41–53. https://doi.org/10.1016/j.coastaleng.2012.04.002

Deza, M., & Deza, E. (2009). *Encyclopedia of distances*. Springer Verlag.

Di Biase, V., Hanssen, R. F., & Vos, S. E. (2021). Sensitivity of near-infrared permanent laser scanning intensity for retrieving soil moisture on a coastal beach: Calibration procedure using in situ data. *Remote Sensing*, *13*(9). https://doi.org/10.3390/rs13091645

Di Biase, V., Kuschnerus, M., & Lindenbergh, R. C. (2022). Permanent laser scanner and synthetic aperture radar data: Correlation characterisation at a sandy beach. *Sensors*, *22*(6). https://doi.org/10.3390/s22062311

Ertöz, L., Steinbach, M., & Kumar, V. (2003). Finding clusters of different sizes, shapes, and densities in noisy, high dimensional data. In *Proceedings of the 2003 SIAM international conference on data mining* (pp. 47–58). Society for Industrial; Applied Mathematics.

Ester, M., Kriegel, H.-P., Sander, J., & Xu, X. (1996). A density-based algorithm for discovering clusters in large spatial databases with noise. *Proceedings of the Second International Conference on Knowledge Discovery and Data Mining*, 6.

GeoTiles.nl. (2021). Ahn: Actueel hoogtebestand nederland [Dataset]. https://GeoTiles.nl

Glennie, C. (2007). Rigorous 3d error analysis of kinematic scanning lidar systems. *Journal of Applied Geodesy*, *1*(3), 147–157. https://doi.org/doi.org/10.1515/jag.2007.017

Grabemann, I., & Weisse, R. (2008). Climate change impact on extreme wave conditions in the North Sea: An ensemble study. *Ocean Dynamics*, *58*(3), 199–212. https://doi.org/10.1007/s10236-008-0141-x

Harmening, C., & Neuner, H. (2020). A spatio-temporal deformation model for laser scanning point clouds. *Journal of Geodesy*, *94*(2), 1–25. https://doi.org/10.1007/s00190-020-01352-0

Hladik, C., & Alber, M. (2012). Accuracy assessment and correction of a lidar-derived salt marsh digital elevation model. *Remote Sensing of Environment*, *121*, 224–235. https://doi.org//10.1016/j.rse.2012.01.018

Hodgson, M. E., & Bresnahan, P. (2004). Accuracy of Airborne Lidar-Derived Elevation. *Photogrammetric Engineering & Remote Sensing*, *70*(3), 331–339. https://doi.org/10.14358/PERS.70.3.331

Holman, R. A., & Stanley, J. (2007). The history and technical capabilities of Argus. *Coastal Engineering*, *54*(6), 477–491. https://doi.org/10.1016/j.coastaleng.2007.01.003

Hulskemper, D., Anders, K., Antolínez, J. Á., Kuschnerus, M., Höfle, B., & Lindenbergh, R. (2022). Characterization of morphological surface activities derived from near-continuous terrestrial lidar time series. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, *48*, 53–60. https://doi.org/10.5194/isprs-archives-XLVIII-2-W2-2022-53-2022

Iglesias, F., & Kastner, W. (2013). Analysis of similarity measures in times series clustering for the discovery of building energy patterns. *Energies*, *6*(2), 579–597. https://doi.org/10.3390/en6020579

IJzendoorn, C. van, Hallin, C., Cohn, N., Reniers, A., & Vries, S. de. (2023). Novel sediment sampling method provides new insights into vertical grain size variability due to marine and aeolian beach processes. *Earth Surface Processes and Landforms*, *48*(4), 782–800. https://doi.org/10.1002/esp.5518

IJzendoorn, C. O. van, Vries, S. de, Hallin, C., & Hesp, P. A. (2021). Sea level rise outpaced by vertical dune toe translation on prograding coasts. *Scientific Reports*, *11*(1), 12792. https://doi.org/10.1038/s41598-021-92150-x

Imparato, D., Teunissen, P., & Tiberius, C. (2019). Minimal detectable and identifiable biases for quality control. *Survey review*, *51*(367), 289–299. https://doi.org/10.1080/00396265.2018.1437947

Intergovernmental Panel on Climate Change (IPCC). (2014). Coastal systems and low-lying areas. In *Climate change 2014 – impacts, adaptation and vulnerability: Part a: Global and sectoral aspects: Working group ii contribution to the ipcc fifth assessment report* (pp. 361–410). Cambridge University Press. https://doi.org/10.1017/CBO9781107415379.010

Jain, A. K. (2010). Data clustering: 50 years beyond k-means. *Pattern Recognition Letters*, *31*(8), 651–666. https://doi.org/10.1016/j.patrec.2009.09.011

Jin, J., Verbeurgt, J., De Sloover, L., Stal, C., Deruyter, G., Montreuil, A.-L., Vos, S., De Maeyer, P., & De Wulf, A. (2021). Monitoring spatiotemporal variation in beach surface moisture using a long-range terrestrial laser scanner. *ISPRS Journal of Photogrammetry and Remote Sensing*, *173*, 195–208. https://doi.org/10.1016/j.isprsjprs.2021.01.011

Jung, Y., Park, H., Du, D.-Z., & Drake, B. L. (2003). A decision criterion for the optimal number of clusters in hierarchical clustering. *Journal of Global Optimization*, *25*(1), 91–111. https://doi.org/10.1023/A:1021394316112

Kellerer-Pirklbauer, A., Bauer, A., & Proske, H. (2005). Terrestrial laser scanning for glacier monitoring: Glaciation changes of the gößnitzkees glacier (schober group, austria) between 2000 and 2004. *Proceedings of the 3rd Symposium of the Hohe Tauern National Park for Research in Protected Areas, Kaprun, Austria*, 97–106.

Keogh, E., & Kasetty, S. (2003). On the need for time series data mining benchmarks: A survey and empirical demonstration. *Data Mining and Knowledge Discovery*, *7*(4), 349–371. https://doi.org/10.1145/775047.77506

Keogh, E., & Ratanamahatana, C. A. (2005). Exact indexing of dynamic time warping. *Knowledge and Information Systems*, *7*(3), 358–386. https://doi.org/10.1007/s10115-004-0154-9

Kerekes, G., & Schwieger, V. (2020). Elementary error model applied to terrestrial laser scanning measurements: Study case arch dam kops. *Mathematics*, *8*, 593. https://doi.org/10.3390/math8040593

Kermarrec, G., Yang, Z., & Czerwonka-Schröder, D. (2022). Classification of terrestrial laser scanner point clouds: A comparison of methods for landslide monitoring from mathematical surface approximation. *Remote Sensing*, *14*(20), 5099. https://doi.org/10.3390/rs14205099

Killick, R., Fearnhead, P., & Eckley, I. (2012). Optimal detection of changepoints with a linear computational cost. *Journal of the American Statistical Association*, *107*(500), 1590–1598. https://doi.org/10.1080/01621459.2012.737745

Koninklijk Nederlands Meteorologisch Instituut. (2022). Uurgegevens van het weer in nederland [website]. https://www.knmi.nl/nederland-nu/klimatologie/uurgegevens

Kuschnerus, M., Lindenbergh, R., Lodder, Q., Brand, E., & Vos, S. (2022). Deteching anthropogenic volume changes in cross sections of a sandy beach with permanent laser scanning. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, *XLIII-B2-2022*, 1055–1061. https://doi.org/10.5194/isprs-archives-XLIII-B2-2022-1055-2022

Kuschnerus, M., Lindenbergh, R., Vos, S., & Hanssen, R. (2024). Statistically assessing vertical change on a sandy beach from permanent laser scanning time series. *ISPRS Open Journal of Photogrammetry and Remote Sensing*, *11*(100055). https://doi.org/10.1016/j.ophoto.2023.100055

Kuschnerus, M., Schröder, D., & Lindenbergh, R. (2021). Environmental influences on the stability of a permanently installed laser scanner. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, *XLIII-B2-2021*, 745–752. https://doi.org/10.5194/isprs-archives-XLIII-B2-2021-745-2021

Kuschnerus, M., de Vries, S., Antolínez, J. A., Vos, S., & Lindenbergh, R. (2024). Identifying topographic changes at the beach using multiple years of permanent laser scanning. *Coastal Engineering*, *193*, 104594. https://doi.org/10.1016/j.coastaleng.2024.104594

Kuschnerus, M., Lindenbergh, R., & Vos, S. (2021). Coastal change patterns from time series clustering of permanent laser scan data. *Earth Surface Dynamics*, *9*(1), 89–103. https://doi.org/10.5194/esurf-9-89-2021

Kuschnerus, M., Lindenbergh, R., & Vries, S. de. (2022). Assessing sandy beach width variations on intertidal time scales using permanent laser scanning. *5th Joint International Symposium on Deformation Monitoring (JISDM)*, 113–119. https://doi.org/10.4995/JISDM2022.2022.13729

Lague, D., Brodu, N., & Leroux, J. (2013). Accurate 3D comparison of complex topography with terrestrial laser scanner: Application to the Rangitikei canyon (N-Z). *ISPRS Journal of Photogrammetry and Remote Sensing*, *82*, 10–26. https://doi.org/10.1016/j.isprsjprs.2013.04.009

Lane, S. N., Westaway, R. M., & Murray Hicks, D. (2003). Estimation of erosion and deposition volumes in a large, gravel-bed, braided river using synoptic remote sens-

ing. *Earth surface processes and landforms: the journal of the British Geomorphological Research Group*, *28*(3), 249–271. https://doi.org/10.1002/esp.483

Lazarus, E. D., & Goldstein, E. B. (2019). Is there a bulldozer in your model? *Journal of Geophysical Research: Earth Surface*, *124*(3), 696–699. https://doi.org/10.1029/2018JF004957

Liao, T. W. (2005). Clustering of time series data - a survey. *Pattern Recognition*, *38*, 1857–1874. https://doi.org/10.1016/j.patcog.2005.01.025

Lichti, D. D. (2007). Error modelling, calibration and analysis of an AM–CW terrestrial laser scanner system. *ISPRS Journal of Photogrammetry and Remote Sensing*, *61*(5), 307–324. https://doi.org/10.1016/j.isprsjprs.2006.10.004

Lindenbergh, R., Kleij, S. van der, Kuschnerus, M., Vos, S., & de Vries, S. (2019). Clustering time series of repeated scan data of sandy beaches. *ISPRS - International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, *XLII-2/W13*, 1039–1046.

Lindenbergh, R. (2010). Engineering applications. In G. Vosselman & H.-G. Maas (Eds.), *Airborne and terrestrial laser scanning* (pp. 237–269). Whittles Publishing.

Lindenbergh, R., & Hanssen, R. (2003). Eolian deformation detection and modeling using airborne laser altimetry. *IGARSS 2003. 2003 IEEE International Geoscience and Remote Sensing Symposium. Proceedings (IEEE Cat. No. 03CH37477)*, *6*, 4113–4116. https://doi.org/10.1109/IGARSS.2003.1295379

Luijendijk, A., & Oudenhoven (Eds.), A. van. (2019). *The sand motor: A nature-based response to climate change: Findings and reflections of the interdisciplinary research program nature coast*. Delft University Publishers - TU Delft Library.

Luijendijk, A., Hagenaars, G., Ranasinghe, R., Baart, F., Donchyts, G., & Aarninkhof, S. (2018). The State of the World's Beaches. *Scientific Reports*, *8*(1), 6641. https://doi.org/10.1038/s41598-018-24630-6

Luijendijk, A. P., Ranasinghe, R., Schipper, M. A. de, Huisman, B. A., Swinkels, C. M., Walstra, D. J., & Stive, M. J. (2017). The initial morphological response of the sand engine: A process-based modelling study. *Coastal engineering*, *119*, 1–14. https://doi.org/10.1016/j.coastaleng.2016.09.005

Maij-Weggen, J. (1990). Kustverdediging na 1990: Beleidskeuze voor de kustlijnzorg. *Ministerie van Verkeer & Waterstaat, 's-Gravenhage*.

Marshall, A. (1972). The thermal properties of concrete. *Building Science*, *7*(3), 167–174. https://doi.org/10.1016/0007-3628(72)90022-9

Masselink, G., & Lazarus, E. D. (2019). Defining coastal resilience. *Water*, *11*(12), 2587. https://doi.org/10.3390/w11122587

Medic, T., Holst, C., & Kuhlmann, H. (2019). Improving the results of terrestrial laser scanner calibration by an optimized calibration process. *Proceedings of the Photogrammetrie, Laserscanning, Optische 3D-Messtechnik, Beiträge der Oldenburger 3D-Tage, Oldenburg, Germany*, 5–7.

Milan, D. J., Heritage, G. L., & Hetherington, D. (2007). Application of a 3D laser scanner in the assessment of erosion and deposition volumes and channel change in a proglacial river. *Earth Surface Processes and Landforms*, *32*(11), 1657–1674. https://doi.org/10.1002/esp.1592

Mukupa, W., Roberts, G. W., Hancock, C. M., & Al-Manasir, K. (2017). A review of the use of terrestrial laser scanning application for change detection and deformation monitoring of structures. *Survey Review, 49*(353), 99–116. https://doi.org/10.1080/00396265.2015.1133039

Muralikrishnan, B. (2021). Performance evaluation of terrestrial laser scanners - a review. *Measurement Science and Technology, 32*(7), 072001. https://doi.org/10.1088/1361-6501/abdae3

Neill, D. B. (2009). Expectation-based scan statistics for monitoring spatial time series data. *International Journal of Forecasting, 25*(3), 498–517. https://doi.org/10.1016/j.ijforecast.2008.12.002

Nicholls, R. J., Wong, P. P., Burkett, V., Codignotto, J., Hay, J., McLean, R., Ragoonaden, S., Woodroffe, C. D., Abuodha, P., Arblaster, J., et al. (2007). Coastal systems and low-lying areas. *Climate change 2007: Impacts, Adaptation and Vulnerability. Contribution of Working Group II to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change*, 315–356.

O'Dea, A., Brodie, K. L., & Hartzell, P. (2019). Continuous coastal monitoring with an automated terrestrial Lidar scanner. *Journal of Marine Science and Engineering, 7*(2), 37. https://doi.org/10.3390/jmse7020037

Park, H.-S., & Jun, C.-H. (2009). A simple and fast algorithm for k-medoids clustering. *Expert Systems with Applications, 36*(2), 3336–3341. https://doi.org/10.1016/j.eswa.2008.01.039

Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., Vanderplas, J., Passos, A., Cournapeau, D., Brucher, M., Perrot, M., & Duchesnay, E. (2011). Scikit-learn: Machine learning in Python. *Journal of Machine Learning Research, 12*, 2825–2830.

Poppema, D. W., Wijnberg, K. M., Mulder, J. P. M., Vos, S. E., & Hulscher, S. J. M. H. (2021). The effect of building geometry on the size of aeolian deposition patterns: Scale model experiments at the beach. *Coastal Engineering, 168*, 103866. https://doi.org/10.1016/j.coastaleng.2021.103866

Puttonen, E., Lehtomäki, M., Litkey, P., Näsi, R., Feng, Z., Liang, X., Wittke, S., Pandžic, M., Hakala, T., Karjalainen, M., & Pfeifer, N. (2019). A clustering framework for monitoring circadian rhythm in structural dynamics in plants from terrestrial laser scanning time series. *Frontiers in Plant Science, 10*. https://doi.org/10.3389/fpls.2019.00486

Ranasinghe, R. (2016). Assessing climate change impacts on open sandy coasts: A review. *Earth-Science Reviews, 160*, 320–332. https://doi.org/10.1016/j.earscirev.2016.07.011

Recuero, L., Wiese, K., Huesca, M., Cicuéndez, V., Litago, J., Tarquis, A. M., & Palacios-Orueta, A. (2019). Fallowing temporal patterns assessment in rainfed agricultural areas based on NDVI time series autocorrelation values. *International Journal of Applied Earth Observation and Geoinformation, 82*, 101890. https://doi.org/10.1016/j.jag.2019.05.023

Rijkswaterstaat - Dutch Ministery of Infrastructure and Water Management. (2022). https://data.overheid.nl/en/dataset/17807-jarkusraaien

Rijkswaterstaat, Rijksinstituut voor Kust en Zee. (2022). Report: Kustlijnkaart (document number 109949). https://open.rijkswaterstaat.nl/open-overheid/onderzoeksrapporten/@133491/kustlijnkaarten-1992/

Rijn, L. C. van. (2022). A fully predictive model for aeolian sand transport, part 3: Verification and application of model for natural beaches. *Coastal Engineering*, *171*, 104051. https://doi.org/10.1016/j.coastaleng.2021.104051

Rofatto, V. F., Matsuoka, M. T., Klein, I., Veronez, M. R., Bonimani, M. L., & Lehmann, R. (2020). A half-century of baarda's concept of reliability: A review, new perspectives, and applications. *Survey Review*, *52*(372), 261–277. https://doi.org/10.1080/00396265.2018.1548118

Schipper, M. A. d., Vries, S. d., Ruessink, G., Zeeuw, R. C. d., Rutten, J., Gelder-Maas, C. v., & Stive, M. J. F. (2016). Initial spreading of a mega feeder nourishment: Observations of the Sand Engine pilot project. *Coastal Engineering*, *111*, 23–38. https://doi.org/10.1016/j.coastaleng.2015.10.011

Schmitz, B., Holst, C., Medic, T., Lichti, D. D., & Kuhlmann, H. (2019). How to efficiently determine the range precision of 3d terrestrial laser scanners. *Sensors*, *19*(6), 1466. https://doi.org/10.3390/s19061466

Schröder, D., & Nowacki, A. (2021). Die atmosphäre als restriktiver einfluss auf messergebnisse eines long range laserscanners. *Proceedings of 21. Internationale Geodätische Woche 2021*.

Schröder, D., Anders, K., Winiwarter, L., & Wujanz, D. (2022). Permanent terrestrial lidar monitoring in mining, natural hazard prevention and infrastructure protection–chances, risks, and challenges: A case study of a rockfall in tyrol, austria. *5th Joint International Symposium on Deformation Monitoring (JISDM)*. https://doi.org/10.4995/JISDM2022.2022.13649

Schubert, E., Sander, J., Ester, M., Kriegel, H. P., & Xu, X. (2017). Dbscan revisited, revisited: Why and how you should (still) use dbscan. *ACM Trans. Database Syst.*, *42*(3). https://doi.org/10.1145/3068335

Schubert, J. E., Gallien, T. W., Majd, M. S., & Sanders, B. F. (2015). Terrestrial laser scanning of anthropogenic beach berm erosion and overtopping. *Journal of Coastal Research*, *31*(1), 47–60. https://doi.org/10.2112/JCOASTRES-D-14-00037.1

Soudarissanane, S. (2016). The geometry of terrestrial laser scanning; identification of errors, modeling and mitigation of scanning geometry [dissertation, Delft University of Technology].

Soudarissanane, S., Lindenbergh, R., Menenti, M., & Teunissen, P. (2011). Scanning geometry: Influencing factor on the quality of terrestrial laser scanning points. *ISPRS Journal of Photogrammetry and Remote Sensing*, *66*(4), 389–399. https://doi.org/10.1016/j.isprsjprs.2011.01.005

Steinbach, M., Klooster, S., Tan, P., Potter, C., Kumar, V., & Torregrosa, A. (2001). Clustering earth science data: Goals, issues and results. *Proceedings SIGMOD KDD Workshop on Temporal Data Mining*, 8.

Stive, M. J., Aarninkhof, S. G., Hamm, L., Hanson, H., Larson, M., Wijnberg, K. M., Nicholls, R. J., & Capobianco, M. (2002). Variability of shore and shoreline evolution. *Coastal Engineering*, *47*(2), 211–235. https://doi.org/10.1016/S0378-3839(02)00126-6

Tan, P., Potter, C., Steinbach, M., Klooster, S., Kumar, V., & Torregrosa, A. (2001). Finding spatio-temporal patterns in earth science data. *Proceedings SIGMOD KDD Workshop on Temporal Data Mining*, 12.

Teunissen, P. (2006). *Testing theory*. VSSD Delft.

Tiberius, C., Marel, H. van der, Reudink, R., & Leijen, F. van. (2021). *Surveying and mapping*. TU Delft Open. https://doi.org/10.5074/T.2021.007

Truong, C., Oudre, L., & Vayatis, N. (2020). Selective review of offline change point detection methods. *Signal Processing, 167*, 107299. https://doi.org/10.1016/j.sigpro.2019.107299

Uphues, C. F. K., IJzendoorn, C. O. van, Hallin, C., Pearson, S. G., Prooijen, B. C. van, Miot da Silva, G., & Vries, S. de. (2022). Coastal aeolian sediment transport in an active bed surface layer: Tracer study and conceptual model. *Earth Surface Processes and Landforms, 47*(13), 3147–3162. https://doi.org/10.1002/esp.5449

van Wiechen, P., de Vries, S., Reniers, A., & Aarninkhof, S. (2023). Dune erosion during storm surges: A review of the observations, physics and modelling of the collision regime. *Coastal Engineering, 186*, 104383. https://doi.org/10.1016/j.coastaleng.2023.104383

Verleysen, M., & François, D. (2005). The curse of dimensionality in data mining and time series prediction. In J. Cabestany, A. Prieto, & F. Sandoval (Eds.), *Computational Intelligence and Bioinspired Systems* (pp. 758–770). Springer.

Vezočnik, R., Ambrožič, T., Sterle, O., Bilban, G., Pfeifer, N., & Stopar, B. (2009). Use of Terrestrial Laser Scanning Technology for Long Term High Precision Deformation Monitoring. *Sensors, 9*(12), 9873–9895. https://doi.org/10.3390/s91209873

Voordendag, A., Goger, B., Klug, C., Prinz, R., Rutzinger, M., & Kaser, G. (2021). Automated and Permanent Long-Range Terrestrial Laser Scanning in a High Mountain Environment: Setup and First Results. *ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences, V-2-2021*, 153–160. https://doi.org/10.5194/isprs-annals-V-2-2021-153-2021

Voordendag, A., Goger, B., Klug, C., Prinz, R., Rutzinger, M., Sauter, T., & Kaser, G. (2023). Uncertainty assessment of a permanent long-range terrestrial laser scanning system for the quantification of snow dynamics on hintereisferner (austria). *Frontiers in Earth Science, 11*. https://doi.org/10.3389/feart.2023.1085416

Vos, S., IJzendoorn, C. van, Lindenbergh, R., & Wulf, A. de. (2024). Asynchronous dune development on a dutch urbanized beach due to buildings and other anthropogenic influences [preprint available at SSRN 4678320]. *Geomorphology*.

Vos, S., Kuschnerus, M., de Vries, S., & Lindenbergh, R. (2023). 4d spatio-temporal laser scan dataset of the beach-dune system in noordwijk, nl. https://doi.org/10.4121/1aac46fb-7900-4d4c-a099-d2ce354811d2.v2

Vos, S., Kuschnerus, M., & Lindenbergh, R. C. (2020). Assessing the error budget for permanent laser scanning on coastal beaches. *Proceedings of FIG Working Week 2020*, (10579).

Vos, S., Lindenbergh, R., & de Vries, S. (2017). CoastScan: Continuous monitoring of coastal change using terrestrial laser scanning. *Proceedings of Coastal Dynamics 2017*, (233), 1518–1528.

Vos, S., Anders, K., Kuschnerus, M., Lindenbergh, R., Höfle, B., Aarninkhof, S., & Vries, S. de. (2022). A high-resolution 4D terrestrial laser scan dataset of the Kijkduin beach-dune system, The Netherlands. *Scientific Data*, *9*(1), 191. https://doi.org/10.1038/s41597-022-01291-9

Vosselman, G., & Maas, H. (2010). *Airborne and Terrestrial Laser Scanning*. Whittles Publishing. https://books.google.nl/books?id=J0DNQQAACAAJ

Vousdoukas, M. I., Ranasinghe, R., Mentaschi, L., Plomaritis, T. A., Athanasiou, P., Luijendijk, A., & Feyen, L. (2020). Sandy coastlines under threat of erosion. *Nature climate change*, *10*(3), 260–263. https://doi.org/10.1038/s41558-020-0697-0

Vriend, H. J. de, Koningsveld, M. van, Aarninkhof, S. G. J., Vries, M. B. de, & Baptist, M. J. (2015). Sustainable hydraulic engineering through building with nature. *Journal of Hydro-environment Research*, *9*(2), 159–171. https://doi.org/10.1016/j.jher.2014.06.004

Weidner, L., Veen, M. van, Lato, M., & Walton, G. (2021). An algorithm for measuring landslide deformation in terrestrial lidar point clouds using trees. *Landslides*, *18*, 3547–3558. https://doi.org/10.1007/s10346-021-01723-4

Wheaton, J. M., Brasington, J., Darby, S. E., & Sear, D. A. (2010). Accounting for uncertainty in DEMs from repeat topographic surveys: Improved sediment budgets. *Earth Surface Processes and Landforms*, *35*(2), 136–156. https://doi.org/10.1002/esp.1886

Wijnberg, K. M. (2002). Environmental controls on decadal morphologic behaviour of the holland coast. *Marine Geology*, *189*(3-4), 227–247.

Williams, J. G., Rosser, N. J., Hardy, R. J., Brain, M. J., & Afana, A. A. (2018). Optimising 4-D surface change detection: An approach for capturing rockfall magnitude–frequency. *Earth Surface Dynamics*, *6*(1), 101–119. https://doi.org/10.5194/esurf-6-101-2018

Winiwarter, L., Anders, K., Czerwonka-Schröder, D., & Höfle, B. (2023). Full four-dimensional change analysis of topographic point cloud time series using kalman filtering. *Earth Surface Dynamics*, *11*(4), 593–613. https://doi.org/10.5194/esurf-11-593-2023

Winiwarter, L., Anders, K., & Höfle, B. (2021). M3C2-EP: Pushing the limits of 3D topographic point cloud change detection by error propagation. *ISPRS Journal of Photogrammetry and Remote Sensing*, *178*, 240–258. https://doi.org/10.1016/j.isprsjprs.2021.06.011

Zhou, K., Lindenbergh, R., & Gorte, B. (2019). Automatic Shadow Detection in Urban Very-High-Resolution Images Using Existing 3D Models for Free Training. *Remote Sensing*, *11*(1), 72. https://doi.org/10.3390/rs11010072

Zimek, A., Schubert, E., & Kriegel, H.-P. (2012). A survey on unsupervised outlier detection in high-dimensional numerical data. *Statistical Analysis and Data Mining: The ASA Data Science Journal*, *5*(5), 363–387. https://doi.org/10.1002/sam.11161

Zoumpekas, T., Puig, A., Salamó, M., García-Sellés, D., Blanco Nuñez, L., & Guinau, M. (2021). An intelligent framework for end-to-end rockfall detection. *International Journal of Intelligent Systems*, *36*(11), 6471–6502. https://doi.org/10.1002/int.22557

ACKNOWLEDGEMENTS

First, I want to give special thanks to Roderik, Sierd and Ramon for their work as my supervisor team, for their guidance, inspiring discussions and encouragement. I am very grateful to Roderik for all the time and patience he took to be my daily supervisor with all the ups and downs this task brings. He has been a great promotor and mentor all these years. Many thanks go to Sander for countless hours of laser scanning, remotely and on the beach, and all the fruitful discussions we had meanwhile. I also want to thank the user committee of the CoastScan project. The regular meetings kept us on track and inspired many new ideas. Thanks to Quirijn and Evelien from Rijkswaterstaat for providing the idea and feedback for part of this research. Thanks to Inés and Valeria for the great collaborations that lead to several publications. I want to thank Katharina for showing me my first point cloud, which convinced me to start this project and for later becoming a role model with her inspiring career. Thanks to Daniel for the great collaboration on many parts of my research and his help on most of the German texts I wrote in the last years, including the summary of this dissertation. A special thank you goes to Adriaan: without his help especially at the beginning of this project, who knows what would have been. And not to forget a big thank you to all my fellow PhD candidates of the Geoscience and Remote Sensing Department. The ones that were there at the start: Antoon, Steven, Inger, Yosra, Mariska, Thore and Malte and those that I met later: Andreas, Felix, Jan, Fritjof, Maaike, Wietske, Sophie, Max, Paco and Daan. I appreciate the supportive community among the PhDs and had a great time.

Finally I would like to thank my family, my parents and my brother and most of all my husband Ricard and our wonderful children. The work on this dissertation has been my job for many years and in some way part of our family life as well. They have been my motivation: their joy and happiness was my compensation for hard work and the reason for persevering.

About the author

Mieke Kuschnerus was born in Gehrden, Germany in 1988 to Dutch-German parents. Her interest for coastal research started at an early age on the Dutch beach of Scheveningen, where she spent many summer days. She completed school at the Siebengebirgsgymnasium in Bad Honnef in 2007. Then she studied Mathematics at the RWTH Aachen and Technical University Berlin, where she completed her Master degree in 2014. Among several jobs as teaching assistant, she spent one semester teaching mathematics to geoscience students at the German University of Oman. There she could make some first observations of dynamics at the sediment surface in the beautiful desert landscapes of Oman.

In 2014 Mieke moved to the Netherlands and started her career at the European Space Agency (ESA) in Noordwijk as a Young Graduate Trainee in the Earth Observation Department. She then continued her work at ESA as mission performance engineer for the Sentinel-6 mission for another three years. She was working on radar altimetry simulator development and estimation of geophysical parameters from raw satellite data of the open ocean. After a training course on radar observations at the Delft University of Technology, she decided to change to research and accepted a position as PhD candidate at the Geoscience and Remote Sensing Department of Delft University of Technology. Her work with a 4D point cloud data set collected with terrestrial laser scanning over a three year period included many trips to the beach in Noordwijk. This dissertation is the result of her research within the CoastScan project over the past five years.