Classification of coastal profile development in the Hoogheemraadschap Hollands Noorderkwartier area

Using advanced data analysis techniques

Nicha Zwarenstein Tutunji

Classification of coastal profile development in the Hoogheemraadschap Hollands Noorderkwartier area

Using advanced data analysis techniques

by

N.C. Zwarenstein Tutunji

to obtain the degree of Master of Science
at the Delft University of Technology,
to be defended publicly on April 8, 2021 at 15:00

Student number:	4618432	
Project duration:	June 8, 2020 – April 8, 2021	
Thesis committee:	Dr. ir. S. de Vries,	TU Delft, supervisor
	Dr. ir. B.C. van Prooijen,	TU Delft
	C.O. van Ijzendoorn Msc,	TU Delft, supervisor
	C. Wegman Msc,	HKV Consultants
	J.K. Leenders Msc,	HKV Consultants
	P. Goessen,	HHNK

This thesis is confidential and cannot be made public until April 8, 2021.

Source cover photo 1 and photo 2

Summary

This thesis describes the development and the characteristics of the coastline along the Dutch coastal zone, managed by Hoogheemraadschap Hollands Noorderkwartier. The coastal zone is a rich environment with multiple stakeholders. These stakeholders use the coastal zone for a range of functions such as safety, nature, business, and recreation. Coastal dunes have been the first line of defense against the sea for many years. On decadal and intercentennial time scale, climate change (e.g. sea level rise) and human intervention (e.g. nourishment), affect the variability of the coastline in different ways. Reports have shown an increase of the mean sea level, which is expected to increase even after the year 2100. Bruun was one of the first researchers, who found a relation between sea level rise and shoreline recession. While Bruuns findings contain the fundamental adaption of the coastline due to sea level rise, the coastline remains a highly complex system.

Morphological data has been collected yearly of the Dutch coast as part of the JarKus program. Meanwhile, the data collection and computational power have increased exponentially, while the computational cost has gone down. Combining the newly computational power with the extensive JarKus data set, provides new insight into the complex coastal system. The derived variables from the JarKus data set range from widths, gradients, volumes to heights. These derived variables are combined with nourishment into a high-dimensional data set.

Clusters of comparable coastal profiles in the Hoogheemraadschap Hollands Noorderkwartier area, have been made using machine learning techniques. Machine learning techniques such as K-means and the Self-Organizing Map (SOM) contain an intelligence, with the capability of clustering similar high-dimensional objects or data, without knowing the desired output. While both methods have the same goal, K-means moves its centroids inside the data and the SOM moves the data to its centroids (BMU). The advantage of combining both methods, is to keep the topological preservation while having the 'hard' clustering advantage of the K-means, which provides easy interpretation and therefore computations.

The coastline of the Noorderkwartier can be broken up into nine clusters. Five of these clusters are classified as main-clusters, having a large number of transects. Four clusters are classified as sub-clusters, having just a few transects. Each of the clusters contains its own characteristic variables. Each of these characteristics originates from its own long-term natural and human-induced causes.

The variable dominating the general clustering, is the active profile of the coastline. The lesser dominating variables are the foreshore nourishment, depth of closure and increase in foreshore volume. Meaning that the high-dimensional data set, find their similarities due to these dominant variables. Upon further investigation, the clusters containing the highest active profiles, were also the clusters containing historical larger nourishments. Comparing the yearly change of the standard deviation, shows that the clusters containing larger historical nourishment, have a shifting depth of closure. With respect to the dunes, correlations are found between the dune foot, y-coordinate of the boundary between marine and aeolian transport, dune volume and active width of the dune. While the exact reason for this correlation is still unknown, it shows potential for further research.

The results of this research contribute to another step in understanding the complex coastal zone. Hoogheemraadschap Hollands Noorderkwartier can now adjust its policy and approach for each cluster based on the results of this research. For further research, it is now possible to focus on specific clusters with their unique characteristics. Lastly, results highlight the importance of the effect of nourishment on the active profile and with this the future dynamic equilibrium of the coastline.

Preface

Before you lies my final work of my Master of Science degree in Hydraulic Engineering at the faculty of Civil Engineering and Geoscience, Delft University of Technology (TU Delft). It presents the results of my research on highly complex coastal systems with the use of advanced data analysis techniques.

It has been a long journey, starting from my HBO civil engineering degree, to finishing my master degree in coastal engineering. It was truly an educational experience, with hard work and lots of fun. The final part, my thesis, has been a very interesting experience with respect to both the subject and my personal growth. Today, I can proudly say that I'm satisfied with the path that I've chosen and can't wait to start my engineering career.

Firstly, I would like to thank my graduation committee for their feedback; Sierd de Vries, Christa van IJzendoorn and Bram van Prooijen. I would like to thank Sierd and Christa in particular, for being my weekly supervisors. With your support I'm able to present this research as it lies in front of you. I'm glad that in these unique Corona times, we were able to have pleasurable weekly meetings online.

Secondly I would like to thank; Carolien Wegman, Jakolien Leenders and Petra Goessen. I would like to thank Carolien, Jakalien and HKV consultants in particular, for allowing me to be part of the team. While having worked from home for most of this research, I've always felt part of the HKV team and enjoyed the weekly team gatherings. I would also like to thank Petra for inviting me on excursions along the coastline of the Hoogheemraadschap Hollands Noorderkwartier area.

Last, but most importantly, I would like to thank my mom (Mylène Zwarenstein). Without your support, I wouldn't have been able to finish my master degree. You've always been an inspiration during my studies and will be so going forward into my engineering career. I would also like to thank my grandmother, who I know looks proudly down upon me.

Nicha Zwarenstein Tutunji
Delft, April 2021

Contents

Nomenclature

Abbreviations	
HHNK	Hoogheemraadschap Hollands Noorderkwartier
IPCC	The Intergovernmental Panel on Climate Change
JarKus	Annual Coastal Measurements
NaN	Not a Number
NAP	Normal Amsterdam Level
PCR	Probabilistic Coastline Recession
RSP	National Beach Pole Line
Self-Organizing Map	
BMU	Best Matching Unit
QE	Quantization Errorr
SOM	Self-Organizing Map
SSE	Sum Of Squared Errors
TE	Topographic Error
U-matrix	Unified Distance Matrix
Variables	
AP	Active Profile
B_ma	Boundary Marine and Aeolian Transport
DF	Dune Foot
DoC	Depth of Closure
DT_prim	Primary Crest
DVol	Dune Volume
FS	Fore Shore
MHW	Mean High Water
MLW	Mean Low Water
MSL	Mean Sea Level

List of Figures

List of Tables

I. Introduction

I.1 Background

Understanding coastal development has been a challenge for engineers throughout the years. In ancient Egypt, the first engineers started exploring the development of river and coastal changes. Many years later, the Vikings made huge ports in Denmark, conquering the sea. Though they still had very little knowledge of the coastal processes, they still managed to see the effects of winds, waves, tides and currents Bruun [1982].

During the Middle Ages, the inhabitants of the Netherlands started building dikes around tidal inlets and rivers, to protect the hinterland against rising water in the rivers. As protection against the rising sea, natural sand barriers were present at the Dutch coast, the dunes Verhagen [1990]. The dunes and dikes made it possible for the Dutch people to inhabit the land, that lies below or just above sea level.

In the years after the Middle Ages, the Dutch dunes became more narrow. Since 1776 Dutch engineers constructed groynes along the coast to trap longshore sediment transport, making the dunes grow Verhagen [1990]. Since 1970, regular nourishment has been part of the Dutch strategy to mitigate coastal erosion.

Hoogheemraadschap Hollands Noorderkwartier (HHNK) is one of the Dutch Waterboards which is continuously working in understanding and managing their coastline for future challenges, with their managed area seen in figure I.1. Morphological data has been collected yearly of the Dutch coast, as part of the JarKus program. This presents the opportunity to use historical data, combined with the increasing popular advanced data analysis techniques Stewart [2019] for new insights in the coastal behavior along the management area of HHNK.

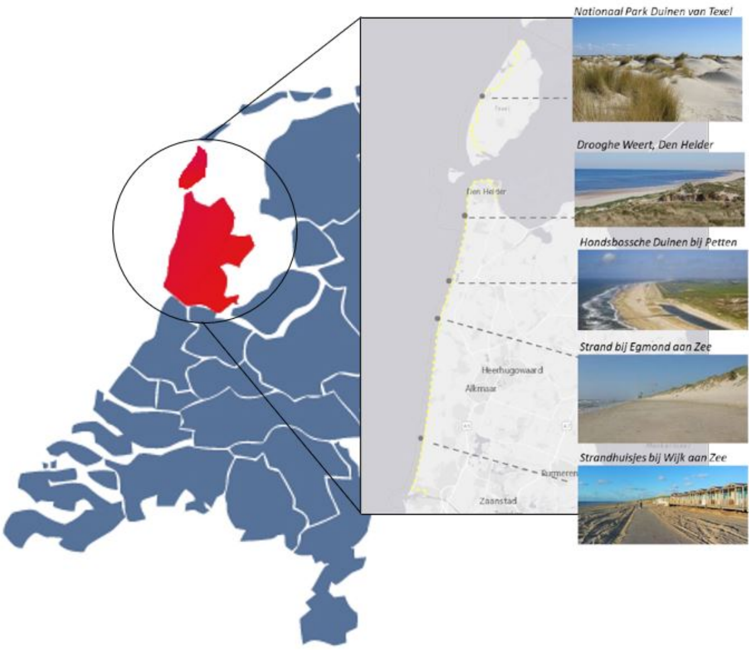


Figure I.1: Management area of Hoogheemraadschap Hollands Noorderkwartier. Wegman and Leenders [2020] Waterschapspartij [2020]

The coastal dunes in the HHNK area, are the first line of defense against the sea. Dunes may form at locations where the wind is able to transport a large supply of sand. The coastline is an ideal place for these processes to occur. At the coast, longshore transport supplies the sand and the waves then transport the sand to the beach. The beach acts as a supply area, where the wind picks up the sand particles in the top layer and transports them to the dunes. The rates of sediment transport are governed by the capacity of the wind, beach dimensions, moisture content, grain size and more De Vries [2019]. The dunes have two main coastal defense functions. The first function is to provide protection to the hinterland. The second function is to provide short term storage of sediment, to allow for adjustment of the beach profile during storms which occur on a small temporal scale Reeve et al. [2002]. Making the dune profile and level of safety dynamic in time.

One of the expected changes of the HHNK coastline in the future, is due to sea level rise. Reports have shown an increase of the mean sea level. It is expected, that the mean sea level will increase even after the year 2100 IPCC et al. [2018]. The Intergovernmental Panel on Climate Change (IPCC) is an independent research institute, that aims to provide transparent and objective research on climate change without biased evaluation. Besides external drivers such as sea level rise, wind, waves, and correlation between local parameters such as the dune volume and the beach slope, also affect the coastline profile De Vries et al. [2012]. Later it was also found, that the beach width is related to dune volume changes Keijsers et al. [2014]. Therefore variables and their changes in space and time can give valuable information about the processes in the complex coastal zone.

Bruun was one of the first researchers, who found a relation between sea level rise and shoreline recession, see figure I.2. Bruun's approach has been criticized by different researchers, due to the fact that Bruun's approach looks at a too confined environment. For example, it doesn't take sediment supply or longshore gradients into account and assumes a static profile shape. Bruun's approach does show a fundamental process, rising sea level acts as a driver of the coast, allowing waves to move up more to the coast and therefore causing more erosion Zhang et al. [2004].

Advanced data analysis techniques have become more popular and have been used in various fields, due to the increase in computational power and data collection Stewart [2019]. Making it easier to apply to historical data derived from the coastline. Advanced data analysis contains an immense branch of methods and algorithms, where finding the right match is a complex task Li [2017]. Understanding the algorithm and results, are the basics in problem solving, with the use of advanced data analytics.

In the last year, research has been done by Van IJzendoorn et al. [2019] to investigate the trend of changes in the Dutch coast, due to the effect of sea level rise. This research has shown that the profile of the Dutch coast does not follow Bruun's rule. The dune toe and crest increase along the Dutch coast, see figure I.3. The governing processes of the increase of elevation, still remain unclear.

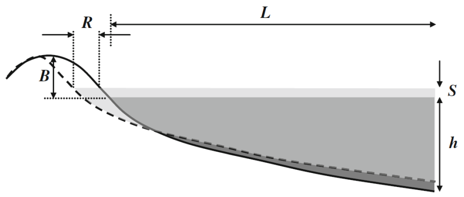


Figure I.2: Schematic diagram showing the Bruun Rule for coastal recession Ranasinghe et al. [2012]

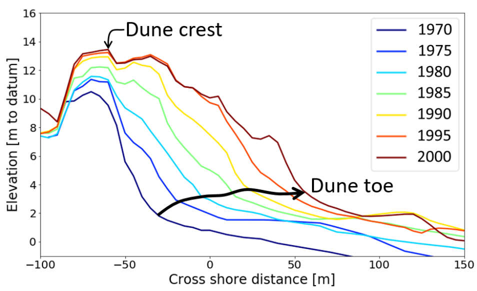


Figure I.3: Elevation of the dune toe and dune crest Van IJzendoorn et al. [2019].

I.2 Purpose of this research

Since the early days, HHNK provides safety against flooding for its managed area. This is done through maintenance and periodic reinforcement of the coast. HHNK works together with multiple stakeholders (e.g. Rijkswaterstaat, the municipality, inhabitants, research and advisory companies such as HKV consultants and businesses) to develop future solutions. These solutions are a balance between safety and spatial adaptation (e.g. drought and water, sustainability, drinking water extraction, nature and recreation). Along the entire Dutch coast, annual measurements of the coastline profiles are done by Rijkswaterstaat Minneboo [1995]. These measurements are abbreviated with JarKus. The goal is to make strategic plans for the future, based on this information. It is not always necessary to intervene, as decisions are made based on risk analyses. Currently the relation between the policy goals and the sediment requirement are not yet clear. With this research, the goal is to give more scientific insight in the complex coastal zone which is managed by HHNK thus future strategical decisions can be made and substantiated.

I.3 Problem description

The Dutch coast and dunes have multiple functions besides safety, such as nature, business and recreation. For future investments, it is important to know how the coastline profiles will change. Coastal engineers have come a long way since the Egyptians and Vikings, but there is still a lot to learn about the complex governing processes at the coast. At this time, predictions on future coastline profile changes of the HHNK area have been made, based on experts opinions and a number of researches. The understanding of the coastal zone still contains many unknown complex physical processes M.R.Hashemia et al. [2010]. In the recent years data collection and computational power has increased exponentially while at the same time the computational cost has gone down. This gives incentive for new data analysis. Measurements such as JarKus, can help validate and provide better understanding of past and future coastline profile changes.

I.4 Research objective

The main objective is to provide deeper and better understanding of the coastal management area of HHNK, based on historical data such as JarKus. The analysis of the historical data is done with the use of advanced data analysis techniques, which have become easier to apply. To do this, the following questions will be answered in this research.

Research question

What is the expected future coastline profile change of the Hoogheemraadschap Hollands Noorderkwartier area based on historical measurements ?

To answer the main research question, a set of sub-questions will be answered.

Sub-questions:

- How can long term changes be characterized along the coastline of the HHNK area, using advanced data analysis and machine learning techniques ?
- Which locations along the coastline of the HHNK area, can be clustered in time and space with their associated characteristics ?
- How do the long term changes along the HHNK area compare with physical understanding of the long term coastal processes ?
- What are the recommendations to incorporate the understanding of long term changes in future strategical plans and policies ?

I.5 Methodology

The main focus of this report is to use historical data to analyse and interpret the complex coastal changes along the HHNK area. Variables are derived from the JarKus data, which are yearly measurements of the coastline profile. Combined with the nourishment data, each transect of the coast contains its own set of variables. These variables represent the unique characteristics of each transect in time and space. This collected data is combined with advanced data clustering techniques: The K-means and the Self-Organizing Map, that uses an intelligence with the capability of clustering similar objects or data without knowing the desired output. As the approach of combining the K-means method and Self-Organizing Map is relatively new, especially with a coastal application, an initial phase was done to understand certain effects in building the data set and model . Finding similarities will give more insight into the relationship between the transect themselves and the relationship between their unique variables. Gaining more insight from an advanced data analysis approach and combining it with the physical understanding. This combination gives new insight into the complex coastal zone managed by HHNK and can be used as input and substantiation for future policy and strategical plans.

I.6 Reading guide

This thesis consists of 5 chapters in total. Chapter I gives an introduction to the purpose of this research and the description of the problem. Chapter II introduces the study area managed by HHNK, the height measurements (JarKus) and changes of the coastline in different spacial and temporal scales. Followed by the geometry of the cross-shore profile of the coastline and its derived variables. Finishing the first part of this chapter with the effect of coastal recession and an explanation of the external drivers. The second part of this chapter starts with a background introduction in the machine learning field of advanced data science. Previous advanced data analysis applications on the coastal zones are researched. On the basis of the problem that is being tackled in this research, two unsupervised machine learning models are introduced, the Self-Organizing Map and the K-means. Chapter III starts with explaining the research phases of the used method, followed by the build-up of the data set. Next, the used hybrid algorithm is explained. Ending this chapter with an explanation of the used methods and data set in both the initial phase and final phase. Chapter IV starts with the results of the initial phase, followed by the results of the final phase. The results of the final phase are used to find and explain the dominant variables found in the clustering. The clusters are referred back to the management area of HHNK. First, the characteristics of the main-clusters are explained in a quantitative and qualitative way. Next the same is done for the sub-clusters. Chapter V is the last chapter of this research. Starting with the discussion, followed by the conclusion and ending this research with the recommendations.

II. Literature study

This paragraph introduces the study area which is management by HHNK (figure II.1) in paragraph II.1. Paragraph II.2 introduces the governing mechanisms on different spacial and temporal scales. Subsequently, the coastline variables are explained in paragraph II.3. Paragraph II.4 looks at the shoreline recession. Paragraph II.5 looks at the external drivers and their relationship to experimental research. Paragraph II.6 introduces the field of advanced data science. Paragraph II.7 gives a detailed explanation of the Self-Organizing Map algorithm. Ending the chapter with a detailed explanation of the K-means algorithm in paragraph II.8.

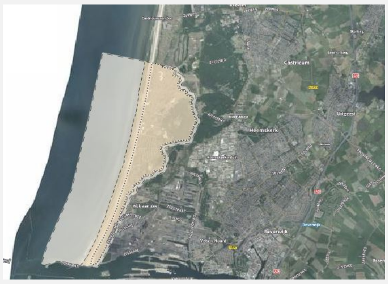
II.1 Study area: Hoogheemraadschap Hollands Noorderkwartier

II.1.1 Management area of Hoogheemraadschap Hollands Noorderkwartier.

.The management area of HHNK consists of a variety of morphological, ecological and economical differences. These varieties can be seen in Figure II.1, where an introduction to the management area of HHNK from north to south is introduced. With Texel being excluded from this research.

Den Helder - Camperduin Underwater at the most northern location, gullies are present along the ebb tidal delta. Along the coast numerous of groynes are present. This location is mostly known for its bathing places, beach houses and beach pavilions. Recently, at the most south location at Camperduin, a large nourishment has been placed. Creating the Hondsbossche Dunes.

Camperduin - Castricum aan Zee Two of the most popular bathing places are located here, Bergen aan Zee and Egmond aan Zee. These two bathing places are widely popular due to the recreational attraction and the Schoorlse Duinen which lie just behind these bathing places. Schoorlse Duinen are the widest dunes found at the Dutch coast, with a great diversity in flora and fauna.



Castricum aan Zee - Velsen-Noord At Castricum aan Zee, a small number of beach pavilions are present. In this area a great number of beach houses are located from north to the south. The area ends at North Sea Channel, with the outer harbor extending far into the sea.

Figure II.1: Introduction to the management area of HHNK (excluding Texel) HHNK [2020].

II.1.2 *Yearly height measurements (JarKus)*

Since 1965 yearly height measurements are taken along transects on the Dutch coast, these are the JarKus measurements. Transects are imaginary lines on water and/or land. Along the Dutch coast, a main transect is located, called National Beach Pole Line (RSP), this line follows the Dutch coast Minneboo [1995]. Every 200 - 250 meters, measurements are done perpendicular to the RSP. Above and around the waterline, these measurements have a spatial resolution of 5 meters. Some meters under the waterline, the spatial resolution is around 10 meters. Early measurements were not accurate and have a larger resolution and lack measurements over time. The quality of the more recent measurements is therefore higher. The measurements are packed in a netCDF file and open sourced by Deltares[1].

The total amount of transects is equal to 2285 transects over the whole Dutch coast for 55 years. Each transect contains the height measurement for each year perpendicular to the coast. The height measurement consists of an x-coordinate, relative to the RSP, and a y-coordinate, relative to Normal Amsterdam Level (NAP). Resulting in a data set that contains 55 years $\times$ 2285 transects = 125675 coastline profiles in space and time.

The management area of HHNK is located in section 6 and section 7 of the JarKus data set. Where section 6 represents Texel and section 7 the area seen in figure II.2. This research will focus on the long term changes in section 7. Section 7 contains 294 transects, with measurements of the last 55 years. This results in 55 years $\times$ 294 transects = 16170 coastline profiles in space and time. Perpendicular to the coast, along the blue line, measurements can be shorter or even not present over 55 years. An example for this can be found in figure II.3, which shows the height measurement of the coastline of the managed area by HHNK for the year 2019.

Figure II.2: Potential measurement coordinates for the management area of HHNK.

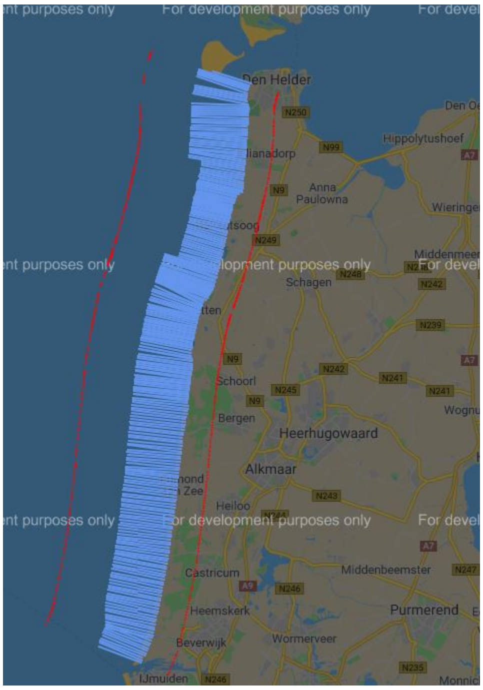


Figure II.3: Actual measurement coordinates (2019) for the management area of HHNK.

[1] http://opendap.deltares.nl/thredds/catalog/opendap/rijkswaterstaat/jarkus/profiles/catalog.html

II.2 Coastline: Spacial scale and temporal scale

The coastline changes due to a variety of natural and human-induced factors. These changes happen on different spatial and temporal scales. This is seen in table 1, which shows the human-induced and natural causes/factors. The understanding of these mechanisms are important for the success of nourishment interventions. The coast shows oscillatory behavior in different temporal scales and can result in nourishment of the coast which would accrete anyway Guille´n et al. [2002]. These mechanisms result in the coastline profile evolving to a dynamic equilibrium under constant forcing for a certain time period.

II.2.1 Inter seasonal and inter annual variability

The inter seasonal and inter annual variability is the smallest temporal scale which is relevant to coastal management. During storms, the upper part of the coastline erodes and accrete when the wave conditions are mild. The Irrabaren number takes into account different forcing parameters and initial conditions of the coastline profile to represent the state in a range of reflective or dissipative equilibrium. If the nourished sediment doesn't differ too much from the native sediment, the natural variability remains almost the same Hamma et al. [2002].

II.2.2 Inter annual and inter decadal variability

Stive et al. [1999] used the JarKus data to look at changes of the coast on decadal scale. The information between the years 1964 and 1992 was used. Two particular areas were investigated, location with human interventions and almost no human interventions. Stive et al. [1999] assumed that the bar system acts as a filter for the wave energy reaching the beach, resulting in a dominant factor on decadal time scale. This was seen in the relation between dune foot change and bar behavior, which travels alongshore on the coast. The research also suggested that cumulative surge-storm parameters correlate with the mean shoreline position Stive et al. [1999].

II.2.3 Intercentennial variability

The JarKus data set contain time scale information in the range of decadal variability to intercentennial variability. This is the main focus of this research for both the available data and future management. As seen in table 1, there are many natural and human-induced causes that affects the coastline on the intercentennial time scale. Understanding the reason behind decadal and intercentennial variability, shore nourishment and natural processes can work in concert rather than counteract each other Hamma et al. [2002].

This research focuses mostly on the intercentennial variability. Combining and understanding all causes and factors is a complicated task. The JarKus data contains the historical variability induced by the different causes and factors. It can be seen in table 1, that relative sea level rise, regional climate variations, coastal inlet cycles and extreme events are natural impacts on intercentennial scale. Phenomenons such as 'sand waves' also impact the changes on the coastline profile. While Stive et al. [1999] observed the dutch coast with the use of the JarKus data set almost 20 years ago, multiple human causes and factors have played a role since then. Mainly nourishment along the coast has increased in recent years.

.

Table 1: Human-induced and natural causes/factors for associated evolutions for shore and shoreline variability (M = Major) Guille´n et al. [2002].

Variability	Time scale\Space scale	Natural causes\factors	Human causes\factors
Late Holocene Variability	centuries to millennia\ 100km+	- 'sediment availability' - relative sea-level changes - differential bottom changes - long-term climate changes - paleomorphology	- human climate change - M river regulation - M coastal structures - M reclamations\closure - structural coastal management
Intercentennial Variability	decades to centuries\ 10 - 100 km	- relative sea-level changes - regional climate variations - coastal inlet cycles - 'sand waves' - extreme events	- river regulation - coastal structures - reclamations\closures - Coastal (non)management - natural resource extraction
(Inter-)annual and decadal Variability	years to decades\ 1 - 5 km	- Wave climate variations - Surf zone bar cycles - Extreme events	- Surf zone structures - Shore nourishment
(Inter-)seasonal and annual Variability	hours to - years\ 10 m - 1 km	- Wave, tide and surges - Seasonal climate variations	- Surf zone structures - Shore nourishment

II.3 Coastline: Variables

II.3.1 Geometry

The cross-shore profile of the coastline can be classified in different zones. In figure II.4, zones in the cross-shore profile are classified by means of derived variables. The cross-shore profile starts with the dune width, which starts from the landward boundary and ends at the dune foot. The location from the dune crest till the dune foot is denoted as the dune front. Close to the dune foot at +2m NAP lies the B_{ma}, which is the border between the marine transport zone and aeolian transport zone De Vries et al. [2010]. From the dune foot till the mean sea level (MSL) is denoted as the beach width Keijsers et al. [2014]. The MSL lies in the middle of the intertidal width, which is the distance between the mean high water (MHW) and the mean low water (MLW). Where often the foreshore width is located from the B_{ma} up to the MLW or taken up until the seaward foreshore location at -4m NAP Van IJzendoorn et al. [2019]. In this research, the foreshore is denoted as the location from the B_{ma} up until the inner depth of closure. This location is often taken at -8m NAP. The outer depth of closure lies around -20m NAP and will be neglected for this research Hinton [2000]. The entire cross-shore profile which is looked at in this research, is the active profile width, which is located from the landward boundary till the inner depth of closure. The visualisation of the variables on the cross-shore profile of the coastal area, can be found in appendix B.

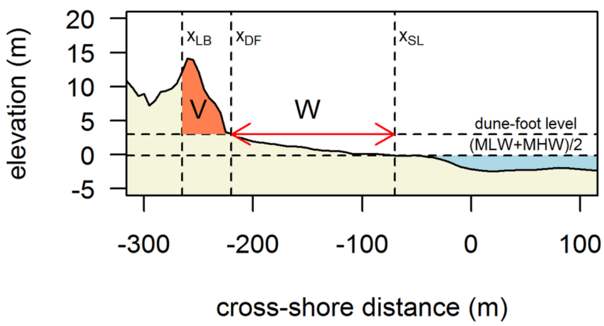


Figure II.4: Definition of different beach zones. Position of the landward boundary (X_{LB}), dune foot (X_{DF}) and the shoreline position (X_{SL}) Keijsers et al. [2014].

Recently Van IJzendoorn et al. [2019] used the JarKus data set to find a correlation between sea level rise and coastal dune changes. Most of the coastline variables as mentioned before in table 2, are derived by Van IJzendoorn using the JarKus data set[2]. By using the x-coordinate and the y-coordinate of the height measurements, variables such as locations, widths, gradients and volumes can be derived along the cross-shore profile of the coast. With the help of these variables, Van IJzendoorn showed that the dune foot and crest increase in elevation along the Dutch coast, not following Bruun's rule.

Table 2: General width definitions of the coastline profile.

Geometry definition	Locations
Dune width	Landward boundary - Dune foot
Dune front	Dune crest - Dune foot
Beach width	Dune foot - MSL
Intertidal width	MHW - MLW
Fore shore width	B_{ma} - Seaward foreshore
Active profile width	Landward boundary - DoC

[2] https://github.com/christa-tudelft/jarkus

II.3.2 Landward boundary

The landward boundary is defined as the landward starting location of the dune. The location has been determined as the most stable landward location from the dune crest with a height greater than 5 meters above NAP and a variance lower than 0.2 Van IJzendoorn et al. [2019] De Vries et al. [2010].

II.3.3 Dune crest

The dune crest is defined as the highest location of the dune. This location can be determined by looking at the prominence. Where a valley is defined as the lowest point between a peak and a higher peak. The prominence is calculated from the peak relative to the highest valley.

II.3.4 Boundary between Marine and Aeolian transport

The coastal area can be classified in a marine transport zone and an aeolian transport zone. The location separating these zones is identified as the boundary, where landward of the boundary no marine processes are present and seaward the changes are mainly due to the marine processes. This location has been assumed to lie around +2m relative to NAP De Vries et al. [2010].

II.3.5 Dune Foot

Multiple researches were done on the exact location of the dune foot. Since 1965, the dune foot was visually determined to be located at about +3 meter relative to NAP Reussink and Jeuken [2002]. This method restricts changes in the vertical. In the research of Van IJzendoorn et al. [2019], two approaches for the dune foot have been used. The first method looks at the first and second derivative of the profile between the MHW and +6 meter relative to NAP. This method looks at the most seaward position, where the first derivative equals 0.001 and the second derivative equals 0.01 Diamantidou et al. [2020]. This method has been found to be a robust estimation of the dune foot. The second method is based on machine learning to predict the location of the dune foot. Beuzen [2019] made his program available called pybeach[3]. This program looks at the location with the highest probability to represent the dune foot location.

II.3.6 Mean High Water and Mean Low Water

Both the MHW and the MLW are defined by taking the average height relative to NAP for a period of 19 years. The MHW has been determined to lie around +1 meter relative to NAP at the Dutch coast and the MLW at around -1 meter relative to NAP Van IJzendoorn et al. [2019].

II.3.7 Mean Sea Level

Defining the mean sea level (MSL) can be done in two ways De Vries et al. [2010]. The first one is to take a fixed vertical position, which is the same position as the Normaal Amsterdams Peil. The second method is to find the average between the MHW and the MLW.

[3] https://pypi.org/project/pybeach/

II.3.8 Depth of Closure - Standard deviation

One of the most important parameters for the Bruun rule, is the depth of closure (DoC). This point is the boundary condition at the seaward side for cross-shore sediment transport. After this point, there is no significant exchange with the beach Aragonés et al. [2017]. The effectiveness of beach nourishment depends on the knowledge of the local DoC, which is often taken at -8 meter relative to NAP De Vries et al. [2010]. It is mentioned by G.Valiente et al. [2019] that there are two ways of calculating the DoC. The first one is through observations and the second one is through numerical models. In figure II.5, the flow diagram can be found for each method. In this flow diagram, it is observed that the dominant forces for the position of the DoC are the wave height and wave period.

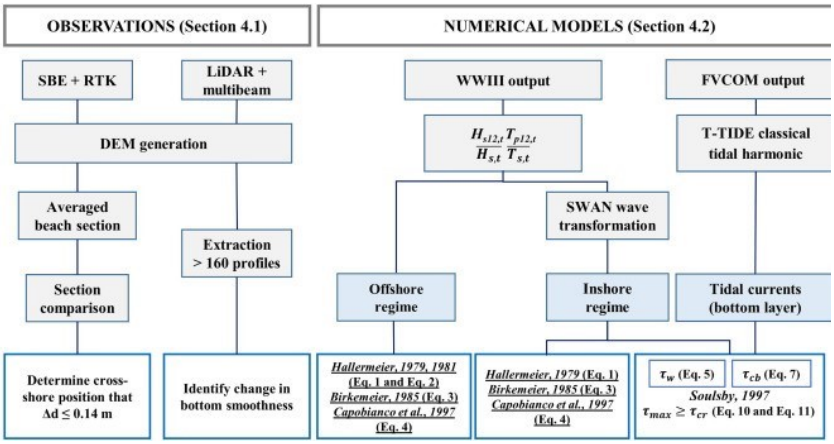


Figure II.5: Flow diagram for defining the depth of closure based on observations and numerical models G.Valiente et al. [2019].

Hinton [2000] researched the DoC at the Dutch coast. Hinton found that the most accurate method for looking at the observations and bottom smoothness, which is seen in figure II.5, the standard deviation of depth change method (ssdc). This method deals with large data sets with outlying values. With the ssdc method, the variation in the standard deviation of elevation as a function of cross-shore distance for x number of profiles along the same alongshore location are located Hinton [2000]. Hinton advises to use a data set of 20+ years. The best method for the Dutch coast, is to find the first value of an ssdc of 0.25 m. Changes are measured every year, at all possible locations in the cross-shore direction, after which the standard deviation is calculated. Figure II.6 shows the ssdc method with a threshold of 0.25 m.

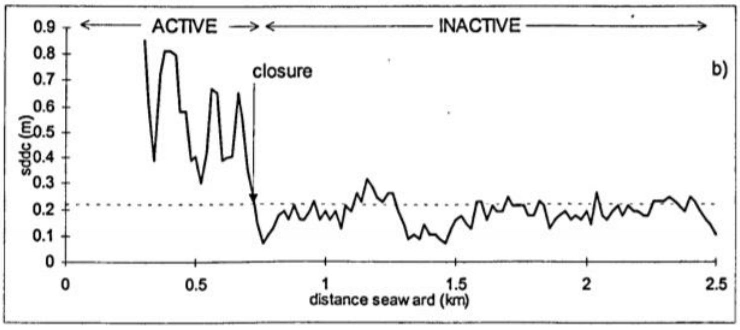


Figure II.6: The standard deviation of depth change method (ssdc) for deriving the depth of closure. With a threshold value of 0.25 m. Depth of closure defines the boundary between morphodynamic active zone and inactive zone Hinton [2000].

II.4 Shoreline recession

II.4.1 *The Bruun Rule*

One of the most famous Danish coastal engineer was Per Bruun. Bruun described that the equilibrium cross-shore profile of the shore shifts upward and landward due to sea level rise Bruun [1962]. It assumes that the cross-shore profile is in equilibrium with hydrodynamic forces such as waves, tides and currents. Due to sea level rise, the cross-shore profile needs to adjust back to its equilibrium. To accommodate for this change in equilibrium, the cross-shore profile is taken as a two-dimensional mass system. Striving for a new equilibrium, the lower part of the cross-shore cross-shore has additional space for sediment deposit. This sediment is then supplied from the upper part of the profile. This effect can be seen in figure II.7. The Bruun rule is expressed in equation II-1.

$$R = \frac{SL}{h+B} = \frac{S}{tan\ \beta} \tag{II-1}$$

Where R is the shoreline recession. S is the sea level rise. L is the horizontal length from the dune peak to the depth of closure. h is the depth of closure, beyond this point no significant sediment transport occurs. B is the height from mean sea level to the dune crest and β is the average slope of the active profile. The schematic positions of these parameters can be found in figure I.2.

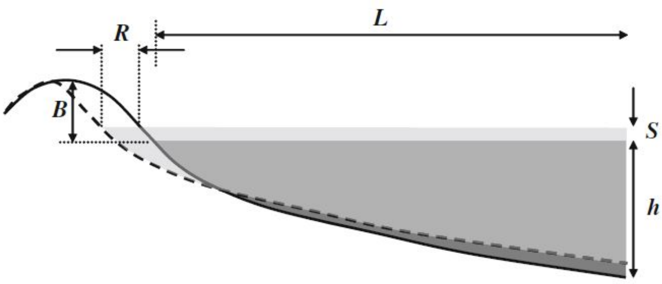


Figure II.7: Schematic diagram showing the Bruun Rule for coastal recession Ranasinghe et al. [2012].

Many researches were done after Bruun, expanding his findings, but at the same time criticizing his confined approach Zhang et al. [2004] Andrew et al. [2004]. Criticizing the Bruun rule as it assumes the cross-shore profile is always in equilibrium, making it hard to apply to coastal areas where the equilibrium changes due to human interventions. Sediment is assumed to only be distributed in the cross-shore direction. Losses due to aeolian transport, gradients in the longshore and transport beyond the inner depth of closure are not taken into account. While the Bruun rule is a highly conservative approach, it does hold fundamental information at the potential consequences of the future coastline.

II.4.2 *Probabilistic Coastline Recession model*

Additionally, the Bruun rule has been expanded with a model that is focused more on the processes which dominate the shoreline recession due to sea level rise. This resulted in a probabilistic coastline recession (PCR) model Ranasinghe et al. [2012]. This model takes into account the probabilistic data of storms over a time span of 10 years and combining it with sea level rise. As the increase of mean sea level rise result in storm waves breaking closer to the coastline. The erosion is then calculated by a simplified wave impact dune erosion model is used Larson et al. [2004]. Hence, the uncertainty in the PCR model is less as it doesn't require highly uncertain estimates of the depth of closure Ranasinghe et al. [2012].

II.5 Coastline: External drivers

II.5.1 Sea Level Rise

Almost 66% of all cities in the world, with a population of five million and more, are located in areas with a risk to sea level rise United Nations [2017]. Predicting the future of sea level rise and its effects, is one of the biggest challenges for engineers nowadays.

The main drivers of sea level rise in the last century are said to be the ocean thermal expansion and the melting of the glaciers IPPC et al. [2013]. Both these effects count for the sum of total sea level rise, see figure II.8. The IPPC follows the Representative Concentration Pathways (RCP), which are different scenarios, based on the changes of anthropogenic behavior. The two scenarios, as seen in figure II.9, are the RCP 2.6 (blue bar) and RCP 8.5 (red bar). Both these scenarios are predicted with a 95th percentile. The trend also shows to increase after 1993. Model based estimations indicate that this is caused by the natural and anthropogenic radiative forcing IPPC et al. [2013].

Besides sea level rise, the ocean thermal expansion and melting glaciers, have effects on the ocean's conveyer belt Pietrzak [2011]. This causes changes in currents, winds and even the Earth's gravitational field. These effects vary around the world, changing the sea level rise per location. Studies have been done on locations that have a great number of measurements. IJmuiden is one of these locations. These measurements show that the sea level rise in the Netherlands lies below that of the global predicted sea level rise trend. The effect of the glaciers melting, is larger on the equator and smaller further away from the equator.

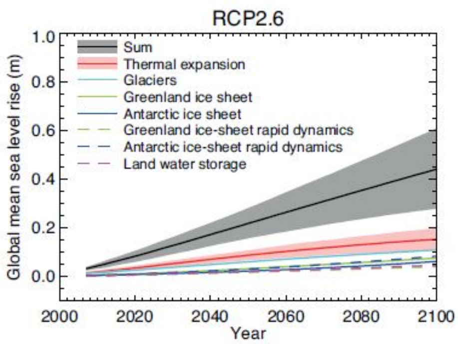


Figure II.8: Projections for global mean sea level rise and their contributions IPPC et al. [2013].

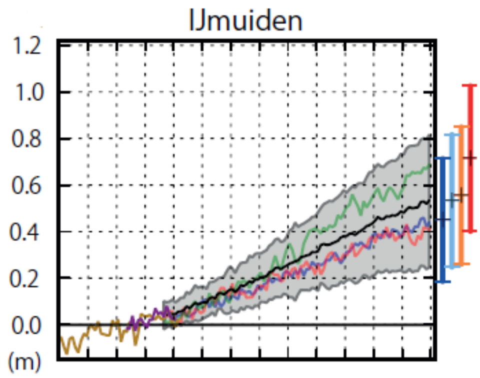


Figure II.9: Projection for global mean sea level rise in IJmuiden, RCP2.6 (dark blue) and RCP8.5 (red) IPPC et al. [2013].

The RCP 2.6 scenario takes into account that CO2 emissions start declining by 2020 and go to zero by 2100. It also takes into account the decrease of different emissions such as methane (CH4) and sulfur dioxide (SO2). These interventions will keep the temperature rise to 2 degrees Celsius and the sea level rise with (0.32 - 0.63) m by the year 2100 IPPC et al. [2013]. The RCP 8.5 scenario is the worst case scenario. This scenario takes into account that the processes are not well understood, and that current emissions are overestimated. The temperature would increase with 4 degrees Celsius and the sea level rise increases by (0.52 to 0.98) m by the year 2100 IPPC et al. [2013]. Hausfather and Peters [2020] states that this case is possible, but the scenario is rather implausible, which is also seen by each passing year. At the same time, KNMI [2017] reported that the findings in IPPC et al. [2013] aren't taking the contribution of Antarctica enough into account. Stating that this will influence the rising of the sea level significantly.

II.5.2 *Nourishment*

Since 1960, Rijkswaterstaat has been providing nourishment along the Dutch coast. As seen in figure II.10, this has been rapidly increasing. Rijkswaterstaat executes the Dutch "Nourishment Program Coastline Care", which aims to prevent structural coastal erosion Deltares [2014]. The criterion for nourishment is the Base Coastline (BKL). The Momentane Coastline (MKL) is a measurement which is done every year. Nourishment aims to prevent the MKL going below the BKL.

Research by Giardino et al. [2019] shows a possible causality between nourishment and the dune foot location. This shows the significance of taking nourishment into account as a space and time variable in the data set. Rijkswaterstaat provided the nourishment data for the entire coast in the years from 1952 up to 2019. Figure II.10 shows how the nourishment in the management area of HHNK, also denoted as section 7. The nourishment in section 7 (blue bar) is compared to the total nourishment along the entire Dutch coast (orange bar).

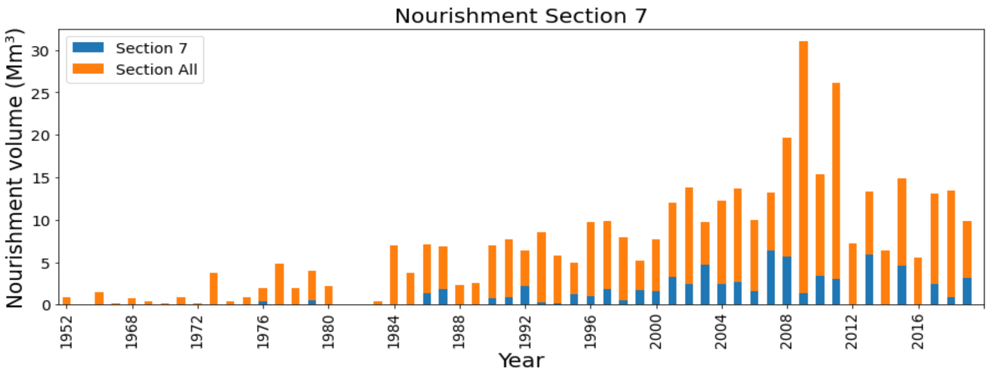


Figure II.10: Nourishment in the management area of HHNK (section 7/blue) compared to the whole dutch coast from 1952 to 2019 (section all/orange).

In figure II.10, an anomaly is removed which represent a large nourishment in 2014. This was the start of the mega nourishment to create the Hondsbossche Dunes. For figure II.11, this mega nourishment is also omitted to scale the graphs. This figure shows that the nourishment is separated into 7 categories. Dune reinforcement was mainly used in the first years. From 1986 beach nourishment was mainly used, up until today. From 1999 beach nourishment has been combined with foreshore nourishment. The nourishment which is done above the MLW, will be combined into **dune nourishment** in this report. All the nourishment done below the MLW will be combined to **foreshore nourishment** in this report.

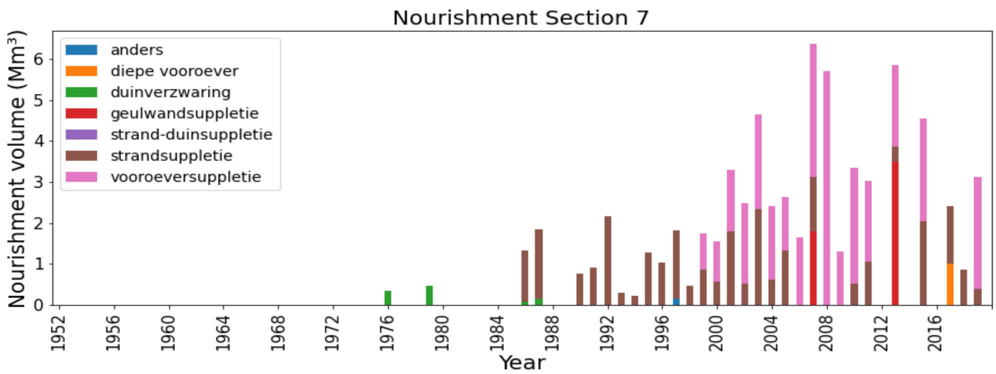


Figure II.11: Types of nourishment in the management area of HHNK (section 7).

II.5.3 *Experiments: Shoreline recession due to sea level rise and nourishment*

Previous research has been done to represent the coastline response of the coast based on sea level rise and nourishment Cowell et al. [1992]. Different responses were modeled with their associated risk for a site specific location. In a more recent report, physical model experiments were done in a flume to look at the effect of sea level rise and nourishment Aktinson and Baldock [2020]. The experimental profile was exposed to a sufficient duration of waves to approach a near equilibrium state. Before the experiments were done, the recession was estimated by expanding the Bruun rule with an additional cross-shore volume term. The experiment consists of four models with different profile types and nourishment locations. Important to note, the nourishment which is reviewed in chapter II.5.2 doesn't contain detailed information about the exact location of nourishment in the profile. One of the four models which represents the foreshore nourishment the most is seen in figure II.12 and figure II.13. This model also came back to represent the expanded Bruun rule the best. Figure II.12 represents the experiment without nourishment and figure II.13 represents the results with nourishment. In both figures, the blue line represents the model before sea level rise and the black line represents the model after sea level rise. The shaded gray area represents the added nourishment.

The recession of the upper part of the profile can be seen in the model with no nourishment. The results represent the Bruun rule, where there is erosion on the upper part of the profile, which is deposited on the lower part of the profile. With nourishment, the erosion on the upper part is minimized, and the lower part of the profile increases in height. The second important result, is that the increase of bottom profile is more spread. Without nourishment, an increase of the bottom profile is seen at a specific location where with nourishment, the increase is spread along with the profile.

As spoken in chapter II.2, both sea level rise and nourishment are important factors on a decadal to intercentennial scale. At the same time, there are many more mechanisms that play a role in the complex dynamics at the coast.

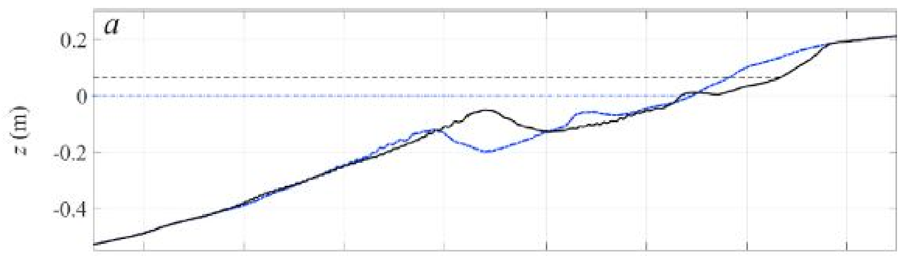


Figure II.12: Coastline profile changes after sea level rise without nourishment [Blue = Profile before sea level rise / Black = Profile after sea level rise] Aktinson and Baldock [2020].

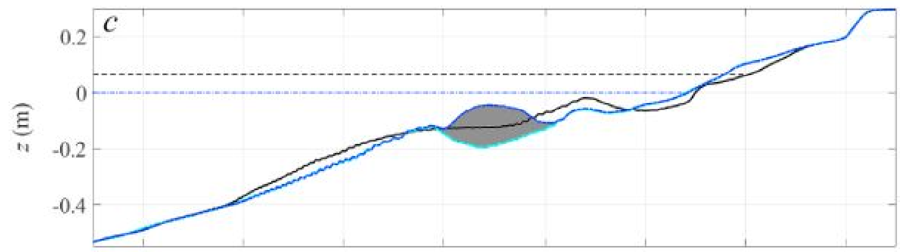


Figure II.13: Coastline profile changes after sea level rise with nourishment [Blue = Profile before sea level rise / Black = Profile after sea level rise] Aktinson and Baldock [2020].

II.6 Advanced data science

II.6.1 Disciplines

Data science is an overarching field for many disciplines. In figure II.14 a few of these disciplines and their mutual coherence are shown. The exact mutual coherence between the disciplines remains a gray area and keeps improving.

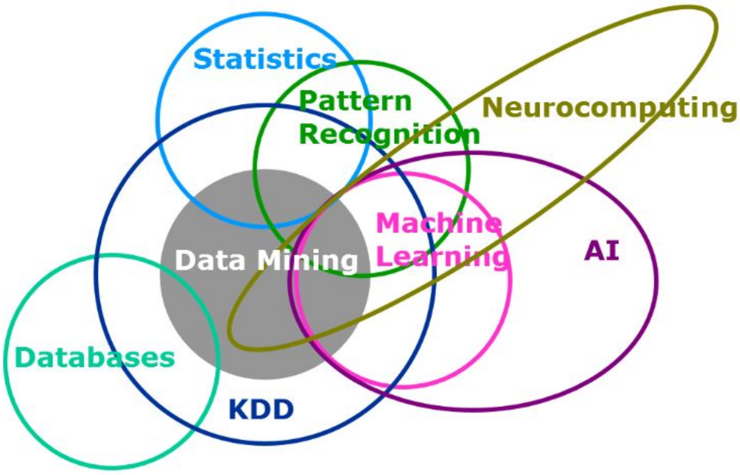


Figure II.14: Venn diagram covering parts of the disciplines in data science Mitchell-Guthrie [2014].

II.6.2 Machine learning vs Statistics

Machine learning is one of the buzzwords in the field of advanced data science Murphy [2018]. While it's becoming more popular, machine learning has been around for a couple of decades. Data collection and computational power have increased exponentially in the last years, while at the same time the computational cost has gone down. Making the use of machine learning more attractive Stewart [2019].

As stated in chapter III.1,lots of information is available from the JarKus and Nourishment data set. Figure II.14 shows a great number of disciplines, that are available to analyse this data. This chapter compares most of the potential data analysis methods to answer the research questions most reliably.

Statistics is a subfield of mathematics, while machine learning is a subfield of computer science and Artificial Intelligence. Statistics rely on mathematical equations, while machine learning requires no prior assumption of the underlying relationships between the variables Stewart [2019]. Matthews analogy about these differences is stated as followed:

> *"Machine learning is all about results, it is like working in a company where your worth is characterized solely by your performance. Whereas, statistical modeling is more about finding relationships between variables and the significance of those relationships, whilst also catering for prediction"*
>
> *- Matthew Stewart*

II.6.3 Machine learning: Algorithms

Machine learning allows computers to learn tricks on their own, without being explicitly programmed. Data is fed to the machine, where an algorithm processes the data and ends in a machine learning model. This model represents what was learned by training the algorithm, for its best performance. The algorithms are classified by four branches for finding its best performance;

Supervised learning: Input data and output data are known. The data is labeled with the output that the algorithm should come up with. The algorithm is trained with training data and compared with the test data. Regression problems and classification problems are the two main areas where supervised learning is often used.
Unsupervised learning: The algorithm doesn't know what the output is. Data is handed unlabeled, not having a correct output or desired outcome. This method is often used in problems such as clustering or anomaly detection.
Reinforcement learning: An agent is placed in an unknown environment. Numerous actions are taken by the agent and is rewarded or punished based on trial and error. The goal is to get the highest reward.
Semi-Supervised learning: There is also a fourth branch, which uses both labeled and unlabeled data.

While the branches narrow down the immense amount of algorithms [4], finding the best matching algorithm remains a complex task Li [2017].

II.6.4 Previous coastal applications using advanced data analytics

Bayesian Approach:
Gutierrez et al. [2012] used the Bayesian network with sea level rising as the main driver. Independent variables included in the Bayesian network consist of beach slope, tidal range, wave height and sea level rise. Resulting in a Bayesian network that provides a probabilistic prediction of the coastline profile change. The accuracy increased with a larger sea level rise per year.

More recently Giardino et al. [2019] applied the Bayesian network to look at the effectiveness of nourishment on the Dutch coast. The JarKus data was combined with historical nourishment data and beach characteristics. Results showed that the current nourishment policy is effective in preventing structural erosion. Effects of future nourishment could be predicted, only if those were seen in the statistical distribution of past nourishment.

Neural Network Approach:
M.R.Hashemia et al. [2010] used a Neural Network to find changes of the coastline profile in the Tremadoc bay (United Kingdom). Twice a year, measurements were taken at 19 stations over a span of 7 years. Input data consisted of significant wave height, significant wave period, angle of the beach, wind direction, and more. The key for the algorithm was, to prevent overfitting, which happens when the model learns to many input and output results. The input vectors were decreased as much as possible and highly correlated data was excluded. Results show that the Neural Network has a great performance in predicting seasonal beach profile changes. M.R.Hashemia et al. [2010] stated that the results generally are more accurate, compared to computationally expensive models for that same region. In that particular research, the data was still very limited.

López et al. [2018] used a Neural Network for sandy beaches, while looking at the effect of marine vegetation. López et al. [2018] states that the accuracy of input nodes isn't as important as finding the right characteristic input nodes. Sixty Neural Networks were made and the best results correlated with wave height, median sediment size, profile slope and an energy reduction factor as input nodes.

[4] https://en.wikipedia.org/wiki/Outline_of_machine_learning

II.7 Data Analysis: Self-Organizing Map

Dr. Eng. Teuvo Kohonen was a Finnish professor and researcher, who became interested in neural networks in 1960. In 1977 he invented the first supervised competitive-learning algorithm, the Learning Subspace Method (LSM). His research on the Self-Organizing Map (SOM) didn't begin until 1981. His idea came from the need for an algorithm, which would map vectors. Vectors that lie close to each other in the input space, are mapped onto contiguous locations in the output space Kohonen [1997]. Self-Organizing Maps are still widely used in data analyses such as the mapping of the Covid-19 virus, spread between countries in 2020 Melin et al. [2020]. Clustering countries behave similarly and thus can benefit by using similar strategies in dealing with the spread of the virus.

II.7.1 Artificial Neural Network: SOM

The Self-Organizing Map (SOM) is a branch of the artificial neural networks (ANNs). It is an unsupervised learning technique that uses a competitive learning technique while maintaining its topological properties from the input space. As seen in figure II.15, the SOM neural network consists of an input layer, weighted matrix and an output layer. The output layer can be 1D, 2D or even a 3D map. For this research, a 2D neuron lattice map is used. These neurons are also called units. The weights in between the input layer and output layer determine the spatial location of each neuron. Over time the weights are trained and updated to cluster the neurons.

The core idea is that the nodes are self organizing, according to their similarity, without any supervision. Thus reducing the high-dimensional data to a lower 2D dimension. Each variable will have its own feature map, also called a component plane, that maintains its topological properties, with respect to the output space.

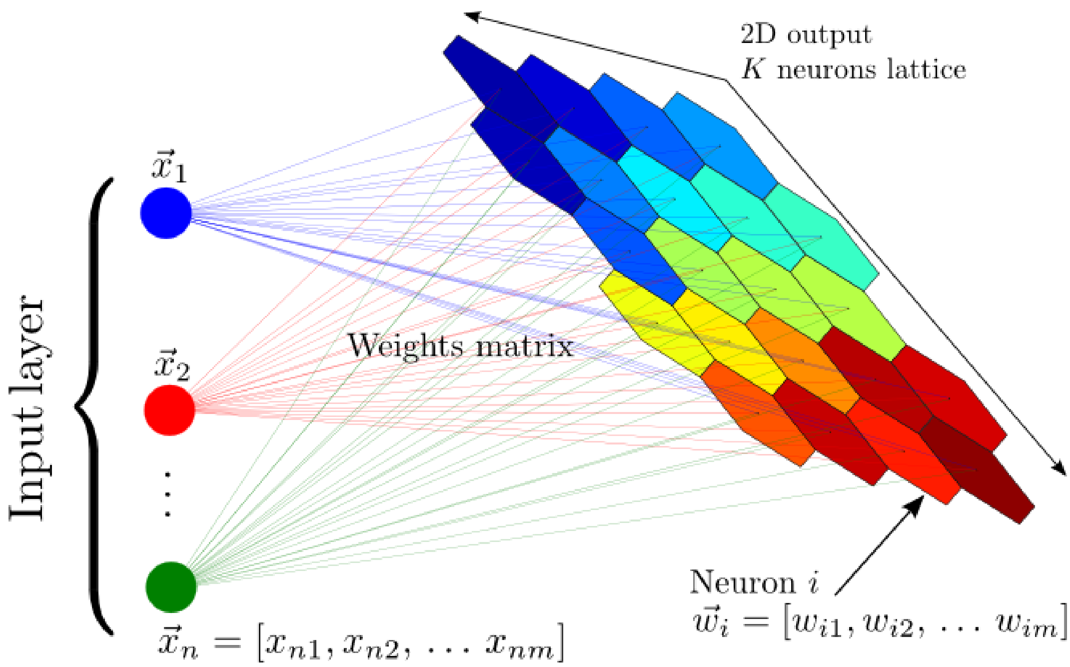


Figure II.15: Self-Organizing Map showing the neural network connections between the input layer and output layer Lan [2018].

II.7.2 SOM algorithm

The SOM algorithm uses competitive learning by adjusting its weights. A single neuron is selected after each iteration, to represent the input vector, while all other neurons compete to represent this vector as well. The winner is called the Best Matching Unit (BMU). The BMU is selected, by calculating the Euclidean difference between all the nodes and the input vector, along with its neighbours within a certain radius. The neighbours positions are slightly adjusted after each iteration to match the input vectors. This is done for all neurons in the output layer before the next iteration. Each iteration is denoted by t as a step. Similar nodes are grouped and dissimilar nodes are separated. Each iteration in the SOM algorithm ends at step 4 and repeats at step 2. The SOM algorithm consists of the following steps Sarkar [2019]:

Step 1. Preprocessing:
The data is normalized and the weights are randomly selected. Additionally the weight can be selected by using the Principal component analysis (PCA).

Step 2. Competition:
Every neuron will compete with other neurons to represent the input vector. This is done by calculating the Euclidean distance (equation II-2), between the weight of each node and the input vector. The node with the smallest Euclidean distance, which is the input vector closest to the weight vector, is called the Best Matching Unit (BMU).

$$||\vec{x} - \vec{w_{ij}}|| = \sqrt{\sum_{t=1}^{n}[\vec{x}(t) - \vec{w}_{ij}(t)]^2} \tag{II-2}$$

With $\vec{x}$ being the input vector and $\vec{w_{ij}}$ being the weight vector connecting the node to the i,j position on the 2D-grid, as seen in figure II.15. The || represents the magnitude of the euclidean distance and t representing the current iteration step.

Step 3. Collaboration:
The topological neighbourhood on the map, starts with the BMU, also the node with the smallest euclidean distance. The radius which affects the neighbours, is represented by a Gaussian distribution. The Gaussian distribution formula representing the smallest euclidean distance can be found in equation II-3.

$$\beta_{ij}(t) = exp(\frac{-min(||\vec{x} - \vec{w_{ij}}||)^2}{2\sigma^2(t)}) \tag{II-3}$$

Where the upper part of the exponential function represents the BMU, which is the node with the smallest distance. The σ value is found in equation II-5. This value represents the radius in which the neighbourhood nodes are affected on the 2D grid.

Step 4. Weight update:
The algorithm learns through adjusting the weights with equation II-4. The neighbourhoods close to the BMU are adjusted to make them more like the BMU. This adjustment is bigger in the beginning and smaller at the end. Also the neighbourhoods that are closer, are adjusted more. This learning is done till the chosen criteria is met.

$$w_{ij}(t+1) = w_{ij}(t) + \alpha_i(t)\beta_{ij}(t)[x(t) - w_{ij}(t)] \tag{II-4}$$

With $w_{ij}(t+1)$ being the updated weight for node i,j. Which contains of the old weight w_{ij} with a slightly adjusted weight to move the node cluster to the BMU. $\alpha_i(t)$ being a factor found in equation II-5, $\beta_{ij}(t)$ being the neighourhood function seen in equation II-3. $[x(t) - w_{ij}(t)]$ being the euclidean distance and t representing the current iteration step

$$\sigma(t) = \sigma_0 \times exp(\frac{-t}{\lambda})$$
$$\alpha(t) = \alpha_0 \times exp(\frac{-t}{\lambda}) \tag{II-5}$$

Both the radius σ and the learning rate α, decrease exponentially over time. With σ_0 and α_0 containing the values of the first iteration and λ representing a constant time value.

II.7.3 Batch Training

To speed up the computation time of the SOM, the batch training mode is chosen in the SOMPY-module Matsushita and Nishio [2010]. This means that the whole data set is trained before it is projected on the map. All vectors are updated according to equation II-6.

$$w_i(t+1) = \frac{\sum_{j=1}^{n} \beta_{ic(j)} \overline{x}_j}{\sum_{j=1}^{n} \beta_{ic(j)}} \tag{II-6}$$

Where $\beta_{ic(j)}$ is the neighbourhood function for the wining node, as seen in equation II-3 and $\overline{x}_j$ representing the mean of the vector.

II.7.4 Output SOM

U-matrix:
The Unified Distance Matrix (U-matrix), as seen in figure II.16, visualises the grouping of x-dimensional input vectors. This is done in a 2D lattice which corresponds with the components plane of each input variable. The clustering can be visualised in a 3D representation of the U-matrix. Similar data is represented by valleys and boundaries are visualised by mountains. Meaning that similar input vectors are located in the same valley.

Components plane:
Each input vector on the U-matrix consist of multiple components planes. Each component or rather variable, is plotted on a component plane while maintaining the same topological location as the U-matrix. An example can be seen in figure II.17.

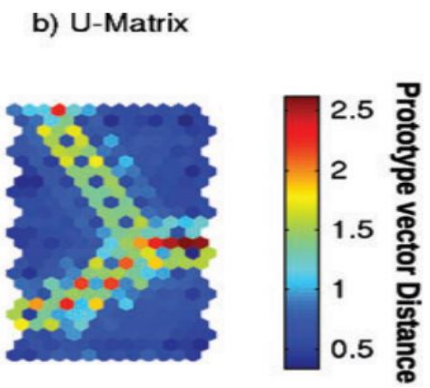


Figure II.16: Example of the unified distance matrix Kohler et al. [2010].

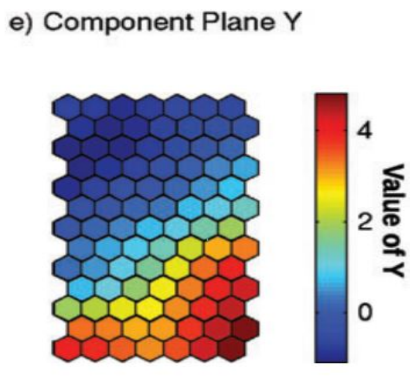


Figure II.17: Example of the component plane Kohler et al. [2010].

II.7.5 *Hyperparameters*

In the previous chapter, parameters such as the weight of the node, are defined based on the results of the model. Hyperparameters require to be defined upfront and control the learning process, as seen in figure II.18. While there are a few hyperparameters that can be tweaked and affect the results, this research will limit itself to the map size and training steps, denoted as variable hyperparameters. The map type, normalization, neighbourhood function and initialization won't be investigated for their effect. These selections are the inputs for the SOMPY module.

Static Hyperparameter

- *Maptype:* A SOM can have different map shapes such as a cylinder and toroid. For this model a static planar hyperparameter is chosen. The nodes can be either rectangular or hexagonal, which influences the number of neighbours. To reduce the difficulty of the map, rectangular nodes are chosen.
- *Normalization:* In the first phase, the data is normalized. This is done by using the unit variance which was divided by the standard deviation.
- *Neighbourhoods:* To define the neighbourhoods from the BMU, a Gaussian distribution is chosen. Both normalization and neighbourhood hyperparameters are commonly used.
- *Initialization:* The initial weights can be random or chosen by utilizing the PCA. With a random sample, the output space will converge differently each time. To increase the automatisation of the algorithm and find the differences between each run, reliability is a key point. Therefore PCA is preferred.

Variable Hyperparameters

- *Mapsize*: Kohonen [1997] suggested to determine the map quality, by looking at and iterating with, the results of the QE and TE. There have been two rules of thumb which have often been used to determine the amount of neurons in the 2D lattice. The first rule of thumb is equal to $Amount\ of\ neurons = 3 \times \sqrt{amount\ of\ input\ vectors}$ Xia [2017]. The second rule of thumb used is equal to $Amount\ of\ neurons = 5 \times \sqrt{amount\ of\ input\ vectors}$ Estévez et al. [2012].
- *Training steps (t)*: The SOM algorithm in chapter II.7.2 is iterated from step 2 to step 4. The number of steps that need to be taken, depends on the converges of the QE and the TE. As these decrease with the number of steps, the quality of the output space will increase.

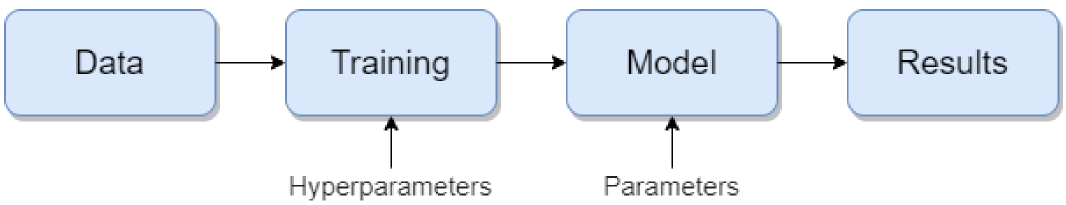


Figure II.18: The associated input in the model for the hyperparameters and parameters.

II.7.6 *Quality evaluation*

Both the Quantization error (QE) and the Topographic error (TE), represent how well the data fits on the two-dimensional map. It also visualises how many iterations are needed to reach convergence. These techniques can be compared to the Sum of Squared Errors (SSE).

Quantization error (QE):
As stated by Kohonen, the best map yields the smallest quantization error (QE) Kohonen [1997]. The quantization error measures the distance between the BMU and the input vectors. The smaller this distance, the closer it lies to its BMU. The formula to calculate this distance is $||x - m_c||$. The average error is calculated for the whole map with equation II-7.

$$QE = \frac{\sum_{i=1}^{N} ||x_i - w_c||}{N} \tag{II-7}$$

Where w_c represents the winning node or the BMU. x_i represents the input vector and N is the number of input vectors, to get the average value.

Topographic error (TE):
The QE only looks at how the nodes, in relation to the neurons, are mapped and not how they are mapped between each other Breard [2017]. For this, the Topographic error (TE) is used. The main feature of the SOM is that the topological relation in a high dimension is preserved in the lower, 2D dimension. The TE measured the discontinuities in the map. The TE is calculated by finding the BMU and the second BMU in the neuron map for each input and evaluated the position. If they are not neighbours, an error is added to t. TE is then equal to the total number of errors divided by the number of data points. This formula can be seen in equation II-8.

$$TE = \frac{1}{n} \sum_{i=1}^{n} t(x_i) \tag{II-8}$$

$$t(x) = \begin{cases} 0 & if\ \mu(x)\ and\ \mu'(x)\ are\ neighbours \\ 1 & if\ not,\ add\ error \end{cases}$$

Where $\mu(x_i)$ is the BMU and $\mu'(x_i)$ is the second BMU.

II.8 Data Analysis: K-means

II.8.1 K-means algorithm

The K-means clustering algorithm was first introduced by MacQueen [1967]. Up until today, K-means is one of the most well-known and used clustering algorithms. It is very easy to implement, is computationally fast and easy to implement on large data sets. K-means is classified as an unsupervised learning technique. The algorithm needs a predefined amount of centroids or K. The number of centroids are randomly placed in and outside of the data. The Euclidean distance between the data is calculated and after each iteration, the centroids is moved to minimize this distance. This distance is the sum of every cluster which contains the sum of the euclidean distance to each data point. This formula can be seen in equation II-9 Tizhoosh [2019].

$$E = \sum_{k=1}^{K} \sum_{x=i}^{X} ||x_i - m_k||^2 \tag{II-9}$$

Where K is the number of clusters, X is the number of data points, x_i the chosen data point and m_k the centroid for cluster k.

An example can be seen in figure II.19. Here a two-dimensional data set is used with three centroids. After iteration, the data set is clustered into three clusters. In this case, the data is separated well and is still very easy to visualize. In a high-dimensional field, k-means can still cluster the data where for humans this is difficult to visualize. If the data isn't separated well, the initial randomly placed centroids can result in different results every time.

II.8.2 Quality control

One of the disadvantages of the K-means is that a predefined amount of centroids or K needs to be selected. Methods such as the "Gap", "Silhouette" and "Elbow" methods are used to identify the possible location of the optimal centroids. This research will focus on using the "Elbow" method as it's often used and the easiest to implement. The Sum of Squared Errors (SSE) decreases with the number of centroids and often the shape of the graph looks like a bent elbow. An example can be seen in figure II.20, where the optimal number of clusters is around three and lies in the pocket of the elbow.

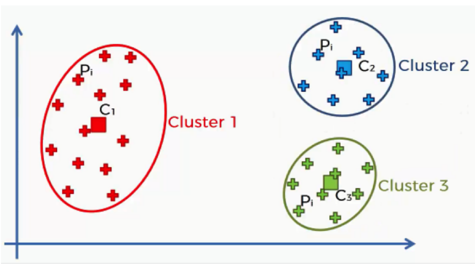


Figure II.19: Example of the K-means clustering with three centroids SuperDataScience [2018].

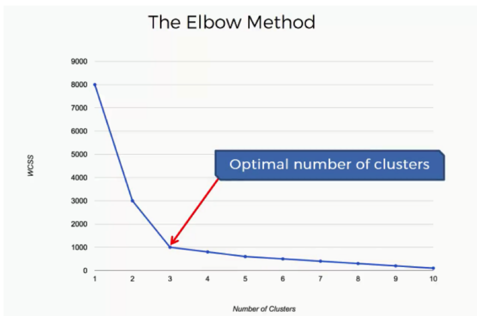


Figure II.20: Example of the K-means "elbow" method SuperDataScience [2018].

III. Methodology

This chapter starts with paragraph III.1, introducing the research phases used in the method. Paragraph III.2 explains the build-up of the data set, followed by the used hybrid algorithm. Ending with steps taken in the initial phase , in paragraph III.4 and the final phase in paragraph III.5.

III.1 Method

III.1.1 Phases

The core component of any machine learning approach, is the input data, from which the algorithm learns. This research will contain not only learning by means of the machine learning algorithm, but also learn from the input data. This approach can be seen in figure III.1. The initial phase, starts with the data set, containing derived variables from the JarKus data and the nourishment data. The origin of the data set is explained more in chapter III.2. Each derived variable in the data set, represents a single individual coastal profile in time and space. The machine learning algorithm, which in this case is a combination of two clustering algorithms, the K-means algorithm and the SOM algorithm. These algorithms learn by repeating and adjusting their algorithm until a certain threshold is reached. After which, the parameters in the final step define the model. The model output contains the clustered input data, with a connection to the input data set. In the initial phase, these results are used to understand the impact of the method of deriving the input variables. In which adjustments are made, resulting in a modified data set. The modified data set is the input for the final phase. Resulting, in the clustering of the final data set, which is used to answer the research questions.

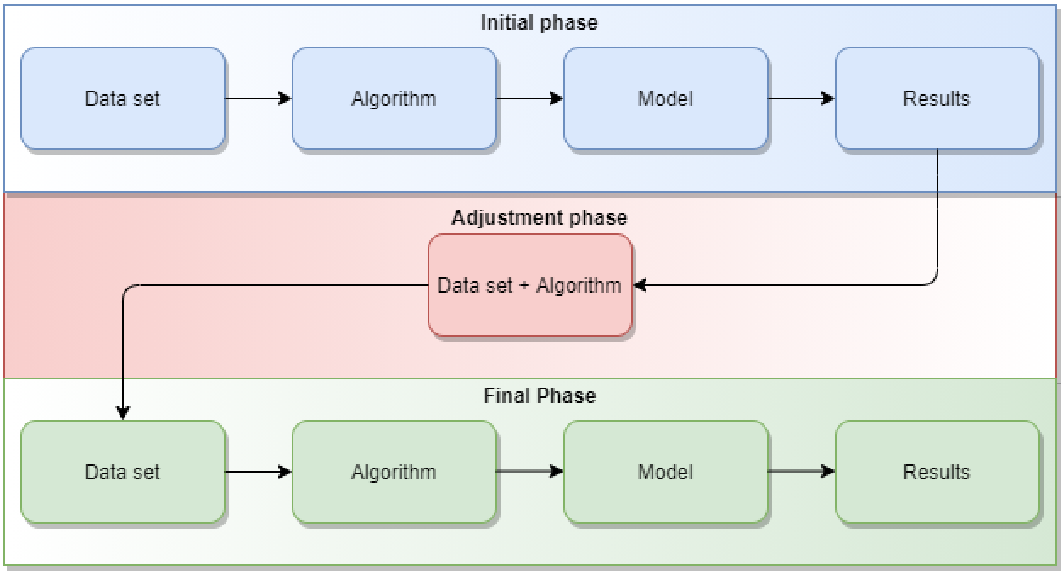


Figure III.1: Method approach

III.2 Data set

III.2.1 Data collection

The source of the data set comes from both the JarKus measurements and the nourishment data. The JarKus measurements are used to derive variables along the cross-shore profile. Close to 50 variables have been derived in previous research Van IJzendoorn et al. [2019]. These variables are extracted for the HHNK management area. These variables are expanded using a revamped standard deviation method, inspired by Hinton [2000]. Resulting in a full data set of 294 transects over 55 years, as seen in figure III.2.

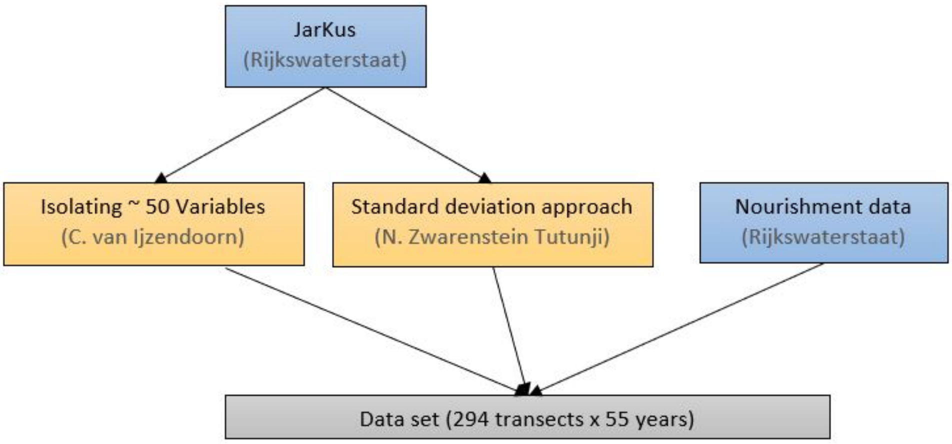


Figure III.2: Sources of data for this research. Blue indicating the raw data, yellow indicating derived variables from the raw data and gray indicating the combined data set.

The cross-shore profile location of the approximated 50 derived variables, can be seen in figure III.3. These variables can be categorized into locations, gradients, widths and volumes. A more detailed explanation of the variables can be found in appendix A and appendix B.

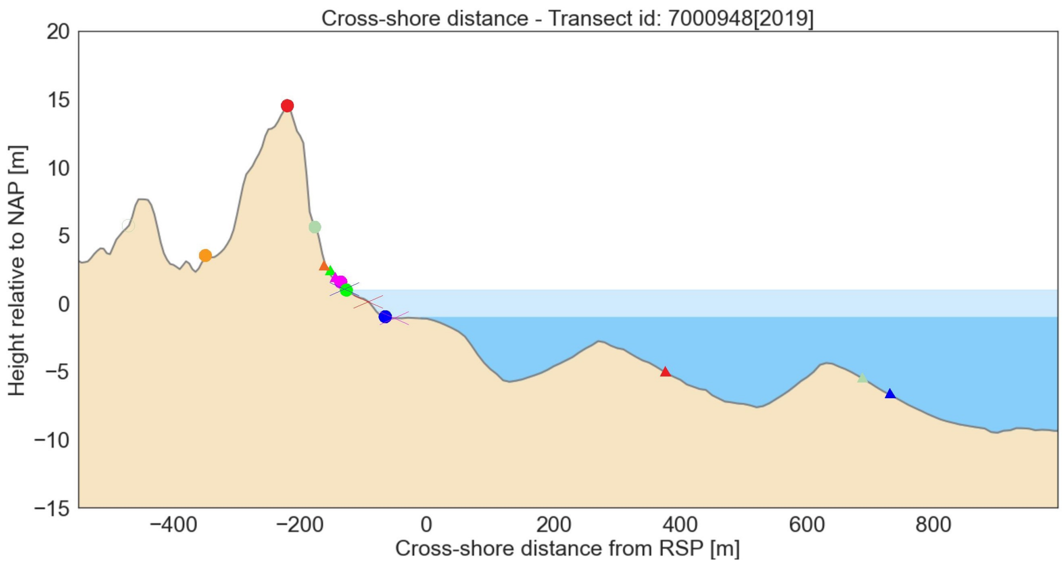


Figure III.3: Locations of the derived variables along the cross-shore profile for transect 7000948.

III.2.2 Data reduction

The height measurements and algorithms used to derive the variables, are not always possible due to certain reasons. These variables result in a 'Not a Number' (NaN) value in the data set. For data preparation, there are several ways to handle NaN values. This can be; omitting the data, assigning zero values, interpolation and extrapolation. This research uses both omitting, interpolation and extrapolation of data to eliminate the NaN values. In figure III.4 all NaN values for each specific derived variable are shown. A couple of variables stand out, which are far above the mean number of NaN values. These specific variables are removed from the data set. After removing the variables with a high number of NaN values, the average amount of missing data rows equal to around 4000 data rows. The total amount of data rows in the data set is equal to 294 transects $\times$ 55 years, which equals 16170 data rows. This means that around 25% of data rows contain NaN values.

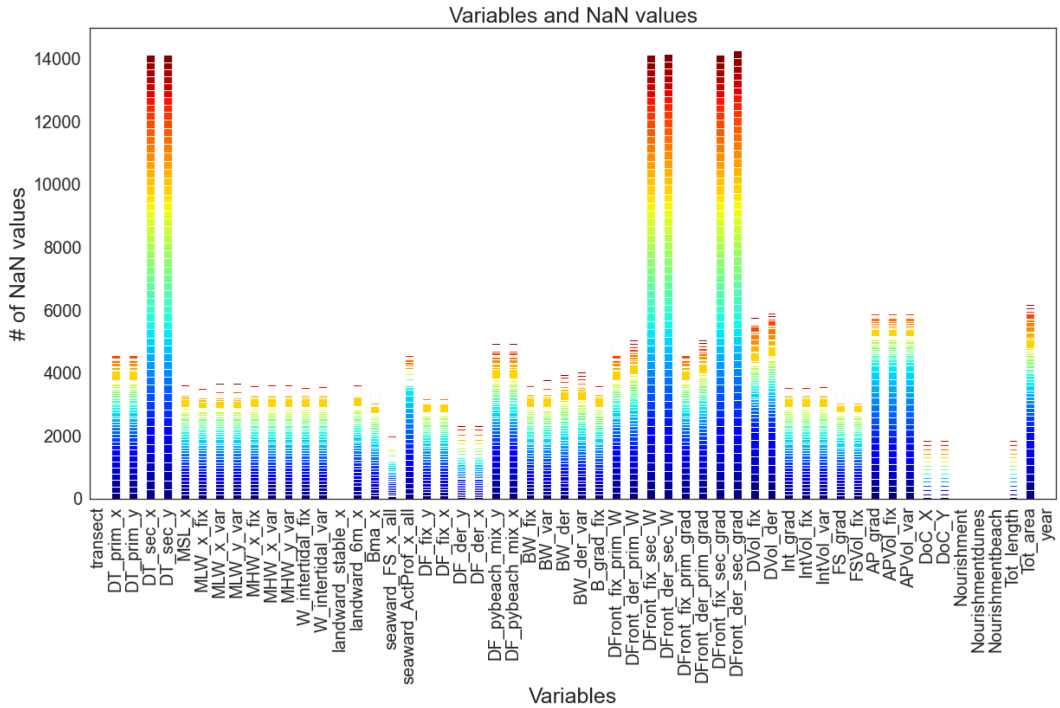


Figure III.4: Number of NaN values for each variable in the data set (Dark red = missing data in 2019 / Dark blue = missing data in 1965).

To investigate the origin of the NaN values, the missing amount of data is displayed in space and time in figure III.5. With dark blue indicating that all data is present and orange indicating that almost all data is missing. A clear pattern of missing data can be seen from before the year 1987, indicated by a light blue area. The northern part of the transects contains a large number of vertical orange bars ranging from the year 1965 to the year 2019. In the horizontal direction, an orange bar can be found in the year 2002, ranging from transect 7000000 up to transect 7005500. The results indicate the locations where the highest quality of data is located.

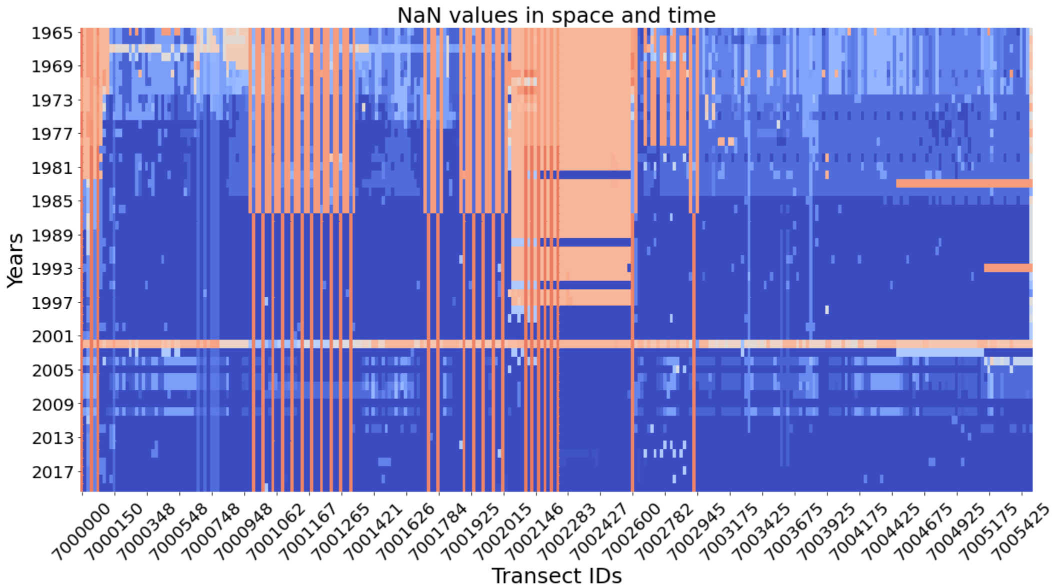


Figure III.5: Amount of NaN values for each variable in the data set.
(Dark blue = 0 number NaN variables / Orange = 40 number NaN variables).

III.2.3 Data format

In the data set, a trade-off is made between having a large enough data set and having the most reliable data. Figure III.5 shows the amount of reliable data in space and time. Based on these results, a separation is made between three areas in space and time. The results of the separated new data set can be found in table 3.

Table 3: Extracted space and time data.

	Area 1	**Area 2**	**Area 3**
Year	1987 - 2019	2001 - 2019	1987 - 2019
Transects	7000000 - 7001925	7001940 - 7002747	7002764 - 7005500

The data set starts out with 25% of missing data, isolating the areas reduces the amount of missing data and increases the reliability. After isolation, transects that contain a vertical orange bar as seen in figure III.5, are omitted. The remainder of the data is interpolated. If interpolation wasn't possible, the remainder transects were omitted as well.

Area 1:
As seen in table 3, area 1 consists of 121 transects over 32 years. After omitting the orange bars from area 1, the data set has an average of 200 NaN values. With 32 years $\times$ 121 transects $=$ 3872 data rows this results in 5.1% of missing data. After interpolation there were still a small number of transects that remained, these were also omitted. Resulting in 93 transects for 32 years.

Area 2:
As seen in table 3, area 2 consists of 58 transects over 18 years. After omitting the orange bars from area 2, the data set has an average of 35 NaN values. With 18 years $\times$ 58 transects $=$ 1044 data rows this results in 3.4% of missing data. All missing data could be interpolated, resulting in 47 transects for 18 years.

Area 3:
As seen in table 3, area 3 consists of 115 transects over 32 years. After omitting the orange bars from area 3, the data set has an average of 250 NaN values. With 32 years $\times$ 115 transects $=$ 3680 data rows this results in 6.7% of missing data. After interpolation there were still a small number of transects that remained, these were also omitted. Resulting in 109 transects for 32 years.

Table 4: Removed transects.

Area 1							
Transects	7000000	7000020	7000030	7000040	7000050	7000060	7000070
	7000150	7000170	7000668	7000708	7000748	7000768	7000789
	7000994	7001023	7001054	7001085	7001115	7001145	7001175
	7001205	7001235	7001265	7001295	7001755	7001784	7001903
Area 2							
Transects	7001932	7001962	7001990	7002015	7002111	7002134	7002158
	7002187	7002212	7002238	7002600			
Area 3							
Transects	7002935	7002987	7003325	7003675	7003800	7005500	

III.3 Machine learning: hybrid algorithm

III.3.1 Clustering

The previous chapter gave an in-depth explanation of machine learning and its different algorithms. Two unsupervised machine learning algorithms were highlighted. The Self-Organizing Map and the K-means method. The goal of both methods is to use an intelligence with the capability of clustering similar objects or data without knowing the desired output. While this makes clustering a very powerful tool in finding structures in data, the results of something unknown are difficult to interpret. Therefore this research will be combined with in depth research in the physical processes based on the results.

III.3.2 Hybrid model (K-means and SOM)

Both the K-means and the SOM model have their advantages and disadvantages. K-means is a "hard clustering", where the data is clustered binary. The SOM is a "soft clustering", where the clustering is identified by likelihood. While the likelihood is a good representation of physical processes, it result in different output each time, where the result can be interpreted differently by each reader.

While fundamentally the methods are the same, the algorithm updates the results based on the lowest Euclidean distance (other methods are also possible). The biggest difference lies in the approach. With K-means the centroids move inside the data to find the cluster and in the SOM, the data moves around the centroids or BMUs. After each iteration, everything is updated in the K-means algorithm, wherein the SOM, only the neurons/units neighbours are updated.

> The Self-Organizing Map can be seen as a constrained K-means.

By combining both models, both the advantages of the K-means and SOM can be utilized. K-means helps the hybrid model to automatise the process and keep the results the same in each run. In the SOM, the results require human intervention to interpret the results. The disadvantage of the K-means clustering is that most of the information in the input space is lost. The SOM maintains the information in the input space due to its topological preservation. The human brain can't cope well with finding patterns in a high-dimensional space and interpret the results. The SOM brings this back to an understandable two-dimensional space.

Combining the topological preservation of the SOM, is done through plotting the contours of the K-means clustering back on the U-matrix. There is still a lot of research in the field of automatisation of the SOM. Lan [2018] expanded the current SOMPY-module by plotting the K-means clusters back onto the U-matrix [5].

While the updated version of the SOMPY-module combines both methods, this research adds a personal addition to the SOMPY-module by adding the contours of the clustering back on each component plane. Thereby finding information from the input space and dominant variable for each cluster.

[5] https://github.com/hhl60492/SOMPY_robust_clustering

III.4 Initial phase

III.4.1 Initial data set

The initial phase starts by using the raw variables derived by Van IJzendoorn et al. [2019]. These variables are derived for each individual profile in space and time. Around 50 variables have been derived and can be categorized into locations, widths, gradients and volumes. Most variables have a fixed position in time, meaning that the variables are taken for the same y-coordinate for each year. Variables are also derived with a variable y-coordinate. The y-coordinate is then calculated through an algorithm or method Van IJzendoorn et al. [2019]. The raw JarKus data is then combined with the nourishment data. This results in around 7000 input vectors in time and space, which are plotted on a 50x60 SOM map. A general overview of the variables in the cross-shore profile can be found in figure III.3. With the more detailed explanation of all variables in appendix A and appendix B.

This initial data set, consists of two data sets, both with different variables. This method is chosen to get more information about the impact of certain variables in the results. The first data set, denoted by 1A, contains the variables which are seen in table 5. The second data set, denoted by 1B, removes the important variables which results from data set 1A. Both models have variables to represents y-coordinates, widths, gradients and volumes. The variables which are based on the x-coordinate, include a biased due to it being relative to RSP. Model 1A takes into account a single x-coordinate, to see the effect of this bias. The second reason for omitted certain data, is due to multiple methods being used in the isolation of the same variable. The fix locations are omitted as much as possible and used if there isn't a variable derived location present. This trade-off results in the chosen data, as seen in table 5.

Table 5: Description of the chosen variables for data set 1A and 1B.

Variable name	**Description**	**Included in data set**
FS_Volfix	Volume between the fixed location in the foreshore	1A
DVol_der	Volume between the variable location in the dunes.	1A
DT_prim_y	Location of the Dune Top	1A + 1B
MHW_y_var	Height relative to NAP of the Mean High Water level	1A + 1B
MLW_y_var	Height relative to NAP of the Mean Low Water level	1A + 1B
DF_der_y	Height relative to NAP of the Dune Foot based on the second derivative approach	1A + 1B
BW_der_var	Beach Width between variable locations	1A + 1B
DFront_der_prim_W	Dune Width between variable locations'	1A + 1B
DFront_der_prim_grad	Gradient between of the dune'	1A + 1B
Int_grad	Gradient of the intertidal location	1A + 1B
FS_grad	Gradient of the foreshore location	1A + 1B
AP_grad	Volume of the active profile, not based on the standard deviation	1A + 1B
DoC_Y	Depth of closure based on the standard deviation	1A + 1B
Foreshore_volume	Volume of the beach, not based on the standard deviation	1B
Dune_volume	Volume of the dunes not based on the Standard deviation	1B

III.5 Final phase

III.5.1 Derivation by means of Standard Deviation

The final phase, replaces the methods for deriving the volumes and locations along the cross-shore profile. This new method uses the standard deviation to derive the locations and volumes along the cross-shore profile. This method originates to identify the depth of closure Hinton [2000]. This research expanse this method to find more locations in the cross-shore profile. With these new variables, extracting of new volumes, widths and gradients are available. Meaning the variables aren't based on fixed locations or different algorithms, but on the same extraction of variability. Figure III.6 shows an example of the standard deviation for transect 7001544. This standard deviation shape characterises most transects along the HHNK area. The landward boundary is found at the red vertical dashed line, the boundary between marine and aeolian transport (Bma_y) at the purple vertical dashed line and the depth of closure at the green vertical dashed line. The volume between the landward boundary and Bma_y is denoted as the new dune volume and the volume between the Bma_y and the depth of closure is denoted as the new foreshore volume.

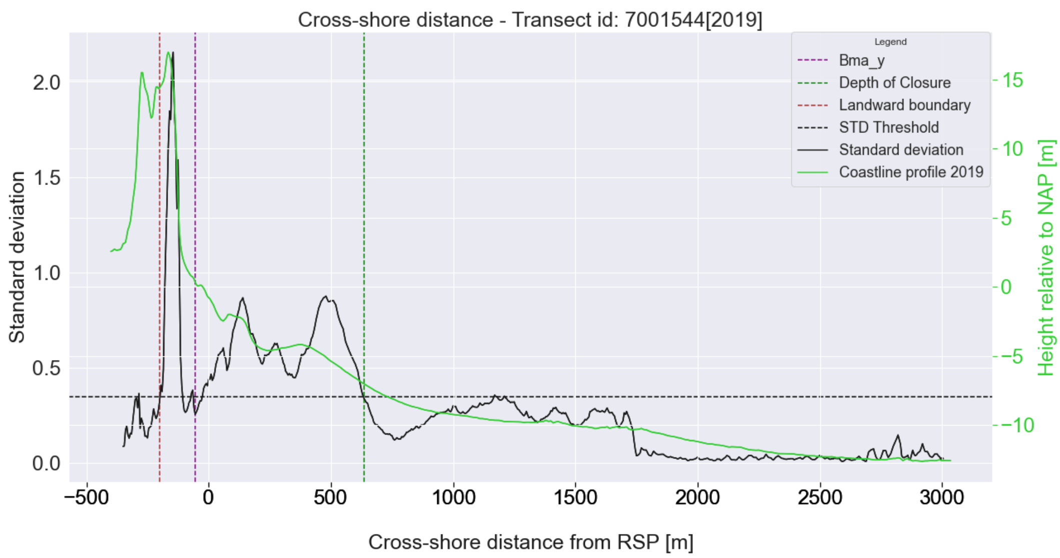


Figure III.6: Derivation of the depth of closure based on the standard deviation for transect 7001544.

III.5.2 Linear regression

The dunes are constantly changing, during a storm, part of the dune is eroded and recovers due to natural processes. Therefore the yearly results of each variable can depend on the yearly amount of storms before a measurement. A regression line, as seen in equation III-1, is plotted in between the data to find a decadal trend. This results in a slope (m) and an average starting value (b). An example for the changes of the dune volume for transect 7000948 in 31 years can be found in figure III.7. In this example the slope m is equal to 5.71 $m^3/year$. The starting volume is equal to 2316 m^3. This method is applied to every transect and variable along the management area of HHNK.

$$y = a \times m + b \tag{III-1}$$

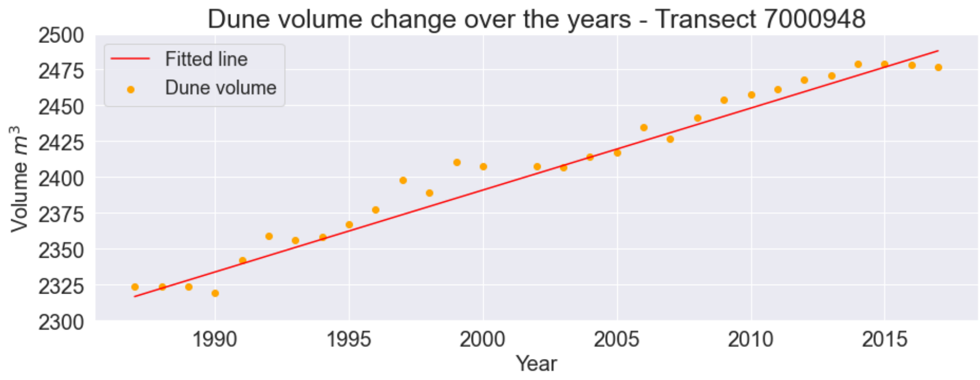


Figure III.7: Linear regression line for the dune volume in transect 7000948.

III.5.3 *Final data set*

In the final data set, the raw JarKus data are expanded and adjusted. The time component of a transect is now represented by a single variable. The total amount of new variables used in the data set, can be seen in table 6.

Table 6: Description of the chosen variables for the final data set.

Variable name	Description
DT_prim_yb	Starting point dune crest
W_intertidal_varb	Starting point intertidal width
BW_varb	Starting point variable beach width
B_grad_fixb	Starting point fix beach width gradient
DFront_der_prim_gradb	Starting point gradient dune front
FS_gradb	Starting point foreshore gradient
DF_y2b	Starting point updated dune foot
beachwidthvar2b	Starting point updated variable beach width
dunefront2b	Starting point updated dune front width
beachy	Derived landward boundary by means of STD
Bma_y	Derived marine and aeolian transport boundary by means of STD
docy	Derived depth of closure by means of STD
Dune_W	Active profile width of the dune by means of STD
ActiveProfile_W	Total active profile by means of STD
Foreshore_W	Active profile width of the foreshore by means of STD
Nourishentforeshoreb	Foreshore nourishment value
Nourishmentdunesb	Dune nourishment value
DT_prim_ym	Slope change of the dune crest height
DT_prim_xm	Slope change of the tune crest horizontal location
W_intertidal_varm	Slope change of the intertidal width
landward_6m_xm	Slope change of the landward 6m location
Bma_xm	Slope change of the marine and aeolian transport border
seaward_FS_x_allm	Slope change of the seaward foreshore horizontal location
DF_x2m'	Slope change of the updated Dune Foot horizontal location
DF_y2m	Slope change of the updated Dune Foot height location
beachwidthvar2m	Slope change of the updated Beach variable Width.
dunefront2m	Slope change of the updated Dune front width.
foreshorem	Slope change of the foreshore volume
dunesm	Slope change of the dune volume

IV. Results

This chapter starts with paragraph IV.1, representing the results of the initial phase. These results are the building blocks to update the algorithm to a final data set. The results of the final data set are represented in paragraph IV.2. The clustered components are represented back on the management area of HHNK in paragraph IV.3. The characteristics of the main-clusters are extensively explained in paragraph IV.4. The characteristics of the sub-clusters are extensively explained in paragraph IV.5. Ending this chapter with deeper research on the variables which dominate the clustering and the mutual correlations IV.6.

IV.1 Initial phase

IV.1.1 U-matrix

The first step utilises the Self-Organizing Map (SOM) algorithm, resulting in the vectors being plotted on the two-dimensional U-matrix. Figure IV.1 represent the U-matrix of data set 1A. Data set 1A shows to have a high number of vectors with a large euclidean distance among each other, both in the middle and in the corners. Comparing these results with data set 1B, there is a more clear distinction between clusters, where the boundary between clusters is separated by a high euclidean distance.

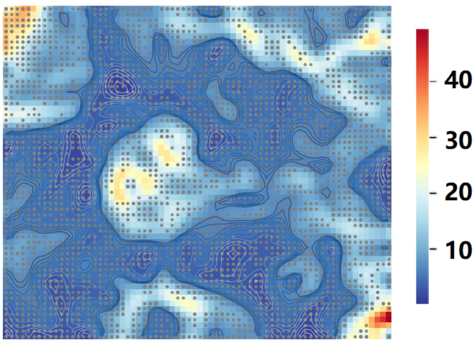


Figure IV.1: U-matrix using the initial data set 1A.

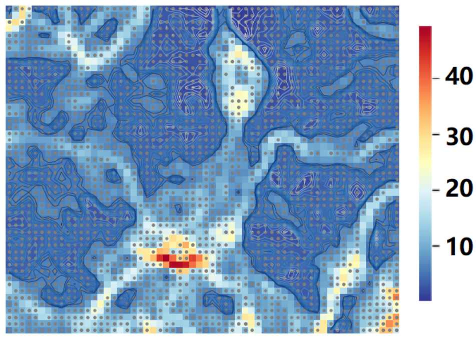


Figure IV.2: U-matrix using the initial data set 1B.

IV.1.2 Clustered U-matrix

As each input vector has a location on the U-matrix, the second step is to cluster each input vector with the K-means algorithm. The results for data set 1A, can be seen in figure IV.3. Here the specific location with high euclidean distance vectors, show to be represented by a single cluster. At the corners we see much larger locations forming a cluster. This shows that the SOM and K-means algorithm deviates from each other. Comparing the results of data set 1A with the results of data set 1B, as seen in figure IV.4, the similarities between the SOM and K-means are more clear. Each enclosed area in the U-matrix, show to be represented by a cluster in the K-means algorithm as well.

Each input vector on the clustered U-matrix, represent an individual profile in time and space. The resulting clusters for each individual space and time profile for data set 1A, can be seen in figure IV.6. The first impression, is that a pattern in both time and space is found. Between the deviation of clusters in time, the clusters are relatively stable. In space, clusters are distinguishable from north to south. Comparing the results of data set 1A with the results of data set 1B, the clusters in figure IV.7 find more similarities in the north and south of the HBD. The south area shows to be dominated by two clusters, which is reflected back in the upper north location. From this north location to the HBD, the area shows to be represented by two clusters.

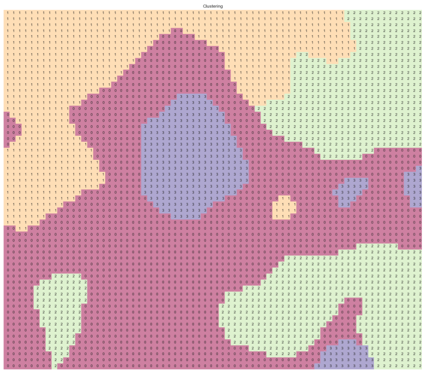


Figure IV.3: K-means clustering on top of the U-matrix using the initial data set 1A.

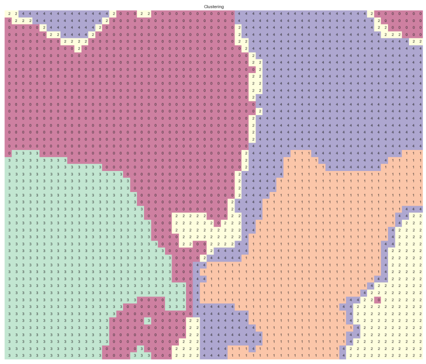


Figure IV.4: K-means clustering on top of the U-matrix using the initial data set 1B.

The contours of the clustered U-matrix can be represented back onto each component plane of the input data set. For the initial data sets, this can be seen in figure IV.5. These individual component planes can be found in appendix C, with the clustered component planes found in appendix D. The clusters in data set 1A show to be dominated by the fixed foreshore volume. The clusters in data set 1B show to be dominated by the volumes and the variable y-coordinate of the MLW and MHW.

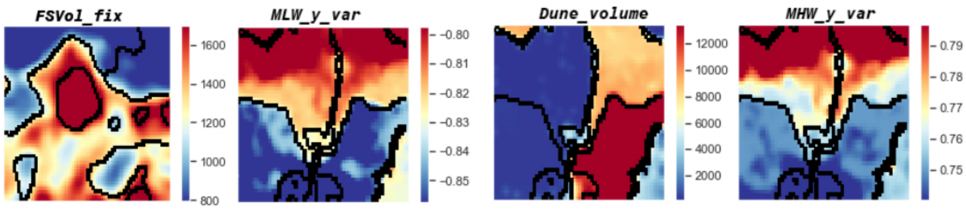


Figure IV.5: clustered component planes using the initial data set 1A and 1B.

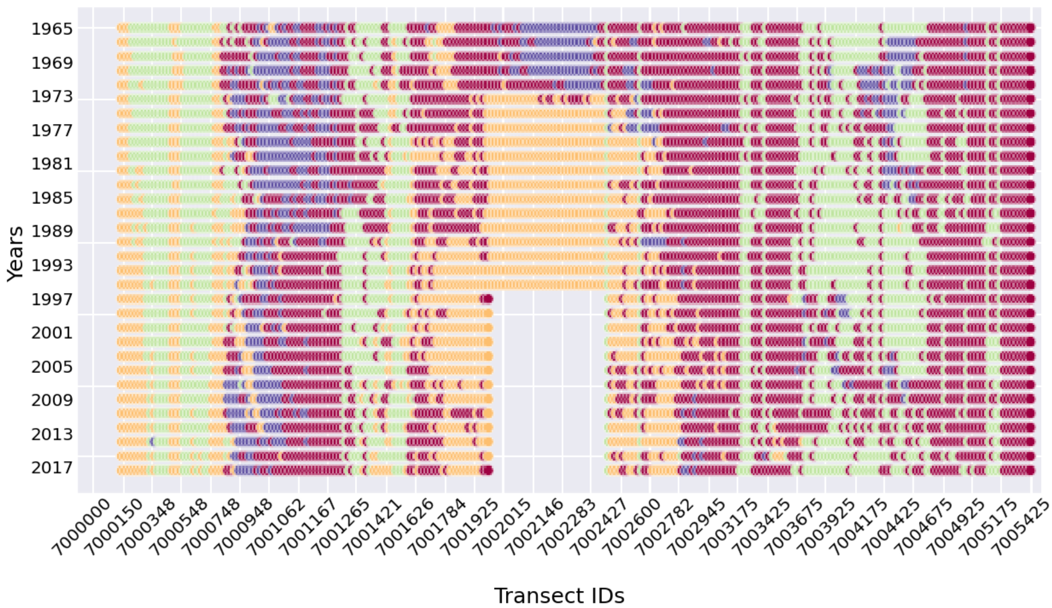


Figure IV.6: Clustered space and time graph using the initial data set 1A.

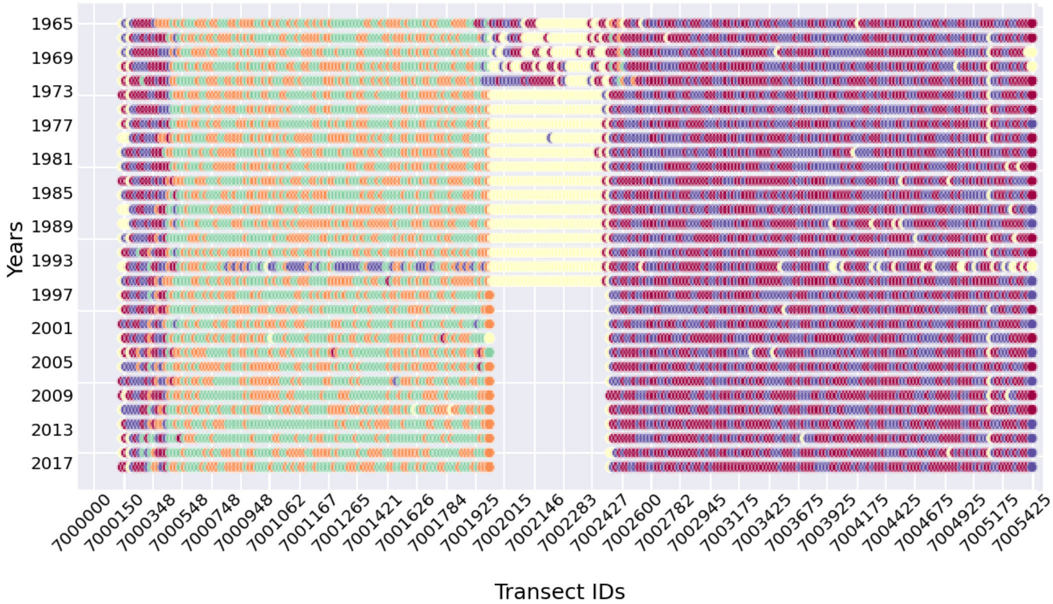


Figure IV.7: Clustered space and time graph using the initial data set 1B.

IV.1.3 *Adjustments*

In the final data set, the extraction of the variables are revamped and are derived based on the standard deviation method. While isolating the depth of closure for initial data set 1A and 1B, a characteristic profile was found, as seen in figure III.6. The results of data set 1A and 1B, show that the volumes play a dominant role in the classification of clusters. In the revamped and final data set, the landward boundary, boundary between marine and aeolian transport and depth of closure are derived with the standard deviation method. This results in new values for the dune volume and the foreshore volume. Beside the volumes, the new locations also results in the active profile width of the dunes and the foreshore. These new variables are the input for the final data set, as seen in table 6.

The results of data set 1A and 1B showed that the clusters are steady in time with a small deviation. The addition of nourishment was hardly reflected in the results. Another improvement that was made, is to incorporate the horizontal movement instead of a horizontal location. This is done with the regression line, as referred in chapter III.5.2. Both the time component, nourishment and horizontal movement are defined by a trend. To identify how these values change from their old profile, the values contain a 'b' base value and a trend value 'm'.

IV.2 Final phase

The final data set contains the adjustments based on the results of the initial data set. These adjustments reduced the input vectors from around 7000 to 250 with an updated derivation of the variables.

IV.2.1 U-matrix

The first step of the final data set is to look at the resulting U-matrix. This result can be seen in figure IV.8. The U-matrix shows to distinguish specific areas in the corners, which are enclosed by high euclidean distances. The middle shows a slight overfitting, where input vectors seem to be clustered individually and not in a group.

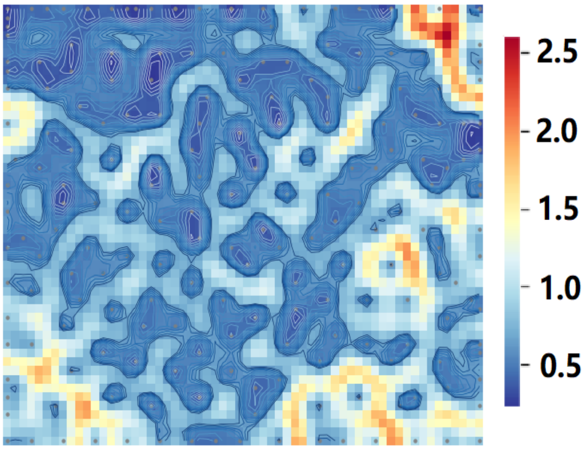


Figure IV.8: U-matrix using the final data set.

IV.2.2 Clustered U-matrix

Repeating what was done in the initial data set, the final data set is also expanded with a K-means clustering. The disadvantage of the K-means clustering, is that the amount of centroids has to be predefined. To further understand the effect of the predefined centroids on the results, the data set is clustered with four, five and six predefined centroids. The results can be seen in figure IV.9. At the most left figure, the U-matrix is clustered with four centroids. The cluster in the lower-left corner is strongly seen back in the U-matrix in figure IV.8. This effect is not seen in the other corners. Therefore the amount of centroids is increased. With five centroids, as seen in the middle figure, the centroid in the lower-right corner is now found. With six centroids, the upper right corner is now also isolated as a cluster, with all the strong clusters following the same contours as the U-matrix. Thereby, using this definition to find the best amount of centroids.

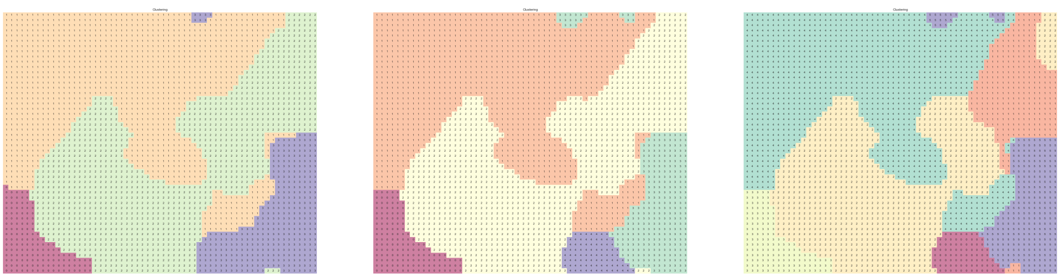


Figure IV.9: K-means clustering on top of the U-matrix using the final data set, with different amount of centroids (k = 4 / k = 5 / k = 6).

The next step is to look at the results of the spatial location for each input vector with respect to the clusters. The results with four centroids can be seen in figure IV.10. The results show, that most of the input vectors which have similarities, are close to each other in the spacial location. Increasing the number of centroids to five, as seen in figure IV.11, the boundaries of each cluster don't change but rather a part of a previous cluster is now defined as an individual cluster. This can be seen at transects 7002882 and 700528. The results of six centroids, as seen in figure IV.12, shows the first sign of cluster which are broken down in the spacial location. This is seen at transect 7001167. This result, combined with the results of the U-matrix, results in the final data set being clustered with 6 centroids.

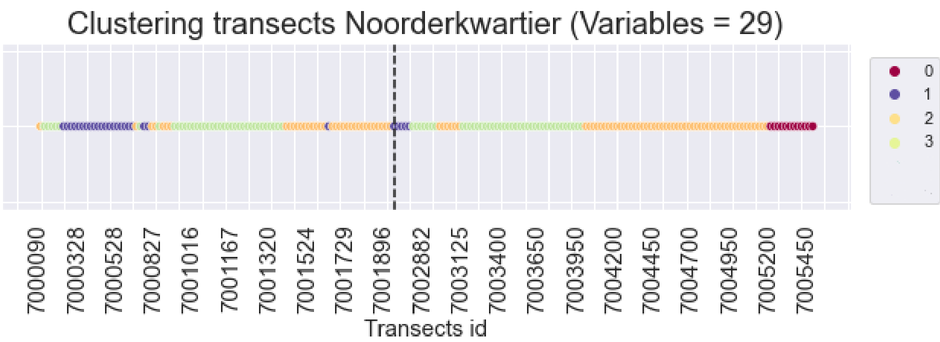


Figure IV.10: Clustered space graph for the final data set with 4 centroids.

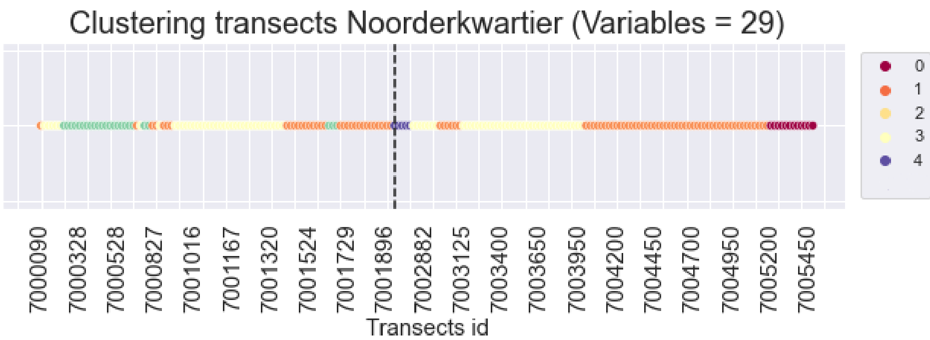


Figure IV.11: Clustered space graph for the final data set with 5 centroids.

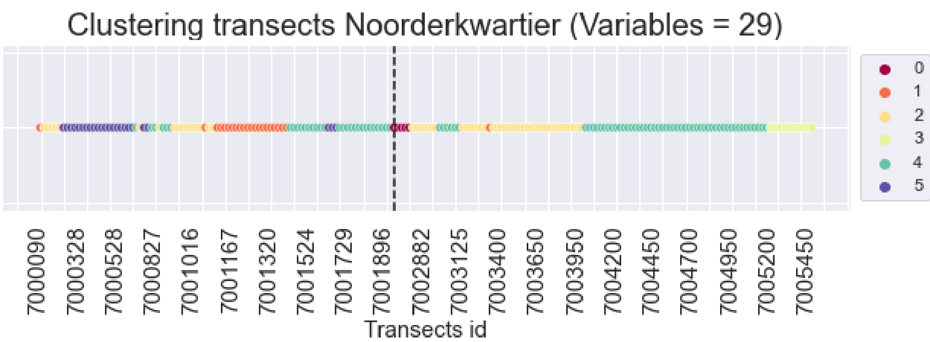


Figure IV.12: Clustered space graph for the final data set with 6 centroids.

IV.2.3 Quality control

As seen in the results, the disadvantage of the K-means algorithm, is predefining the number of centroids. The quality algorithm used for the K-means, is the Sum of Squared Errors (SSE), which is seen in figure IV.13. Typically in the SSE an elbow can be found to justify the number of centroids needed. Looking at the results, the elbow is hardly found in the results of the final data set. The elbow method is a method to find a threshold where adding a centroid doesn't affect a large part of the data set. Meaning, that adding centroids in the final data set, has a small effect. Therefore the number of centroids are taken with respect to the changes in clustered U-matrix.

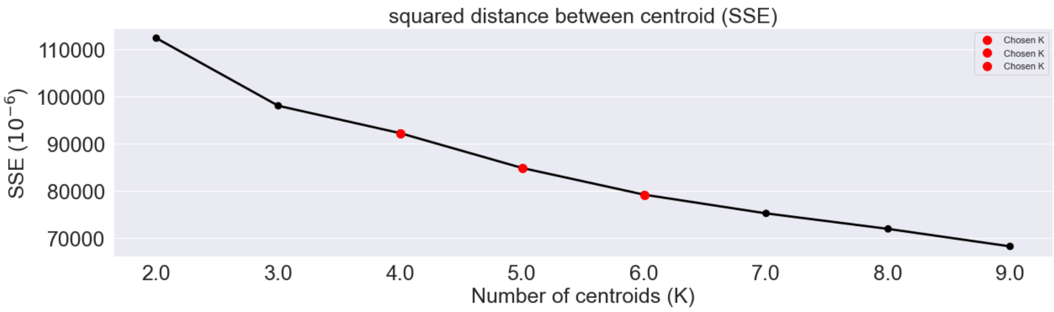


Figure IV.13: Sum of Squared Errors for the final data set.

The quality of the SOM algorithm is measured by both the quantization error and topographic error. The results of both errors can be found in figure IV.14. After around 10 iterations, the quantization error remains stable at around 2.05. The topographic error is very small and therefore negligible in the final data set. The speedup of the SOM by means of PCA can be seen in the results, where the quantization error reaches stability very fast.

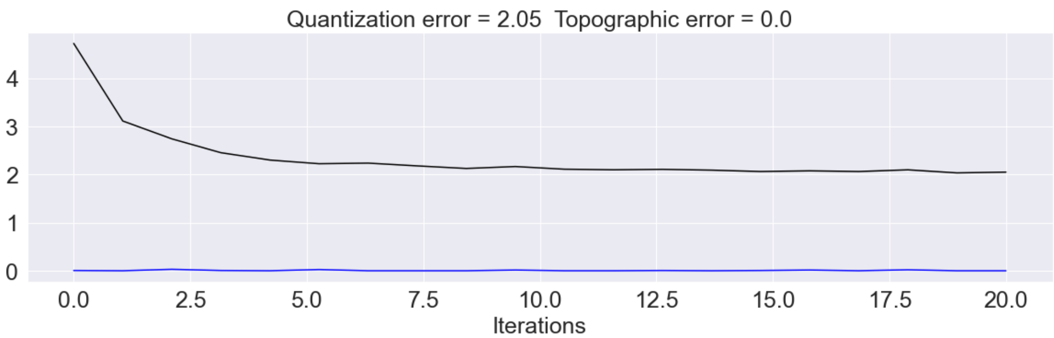


Figure IV.14: Quantization error and Topographic error for the final data set.

IV.3 HHNK clusters

The final data set results in nine clusters along the transects located in the coastal management area of HHNK. These clusters can be identified in the clustered U-matrix in figure IV.15. Each individual cluster holds its own characteristics, based on the combination of known and unknown mechanisms. Clusters having the same color, indicate that the K-means algorithm clustered them together, but in the SOM, these clusters lie separated from each other. As the input vectors in the SOM lie close to similar input vectors, the representation of areas with the same color which don't lie close to each other, will be seen as individual clusters.

The five largest areas containing the most input vectors, are indicated as main-clusters. The remaining four clusters, with a small number of input vectors are indicated as sub-clusters. The location of both the main-clusters and sub-clusters on the U-matrix, can be seen in figure IV.15. Each input vector in the clustered U-matrix are represented back onto the spacial location in figure IV.16.

The spacial representation of the clusters, which are located in the area managed by HHNK. The real-world locations of the clusters north of the HBD, can be seen in figure IV.17. The results south of the HBD can be seen in figure IV.18. The main-clusters are indicated by a solid line and the sub-clusters are indicated as a dashed line. In total 202 transects are clustered. The transects which were omitted due to various reasons as mentioned in this report, are indicated by a gray solid line.

Intermezzo: *The reader is advised to read appendix B to understand the meaning of each used variable and its context.*

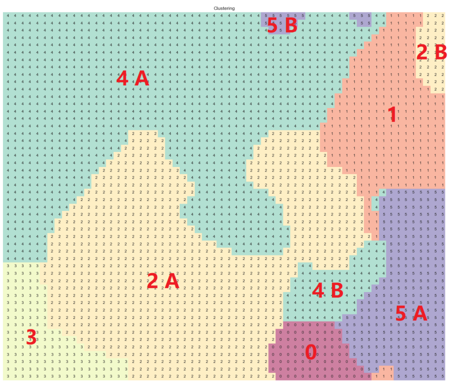


Figure IV.15: Classification of the clustered U-matrix for the final data set. With a total of nine clusters being identified.

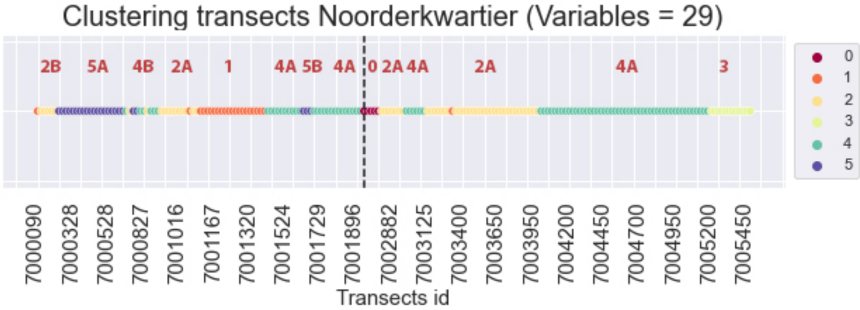


Figure IV.16: Representation of the nine clusters in the space graph for the final data set.

An interactive Google Map webpage of figure IV.17 *and figure* IV.18 *can be found in this* **hyperlink** [6].

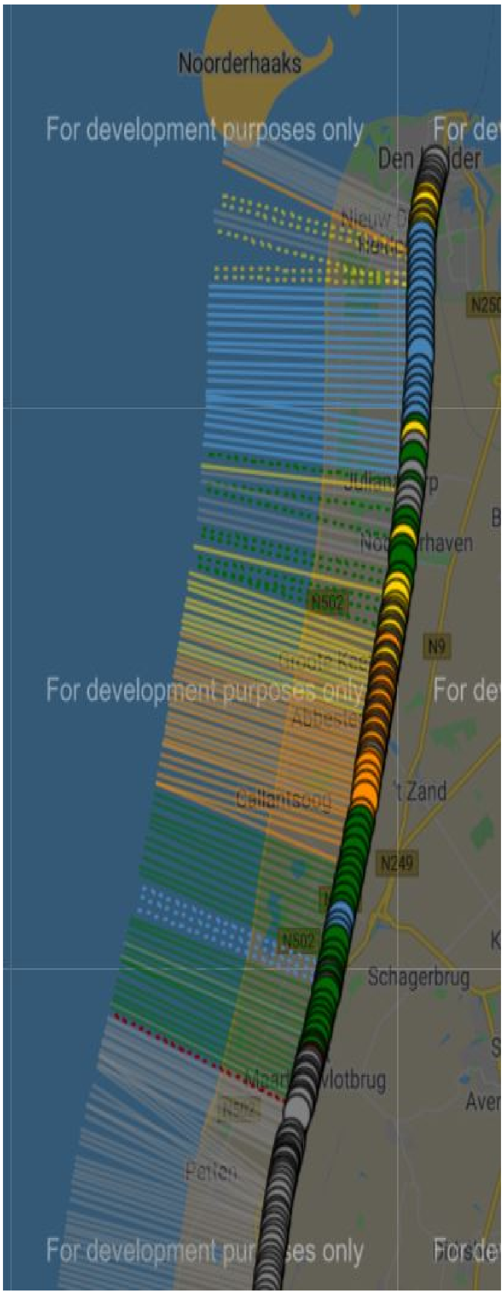


Figure IV.17: Graphical overview of the clusters in the northern part of the management area of HHNK. Solid line representing the main-clusters and the dashed line representing the sub-clusters.

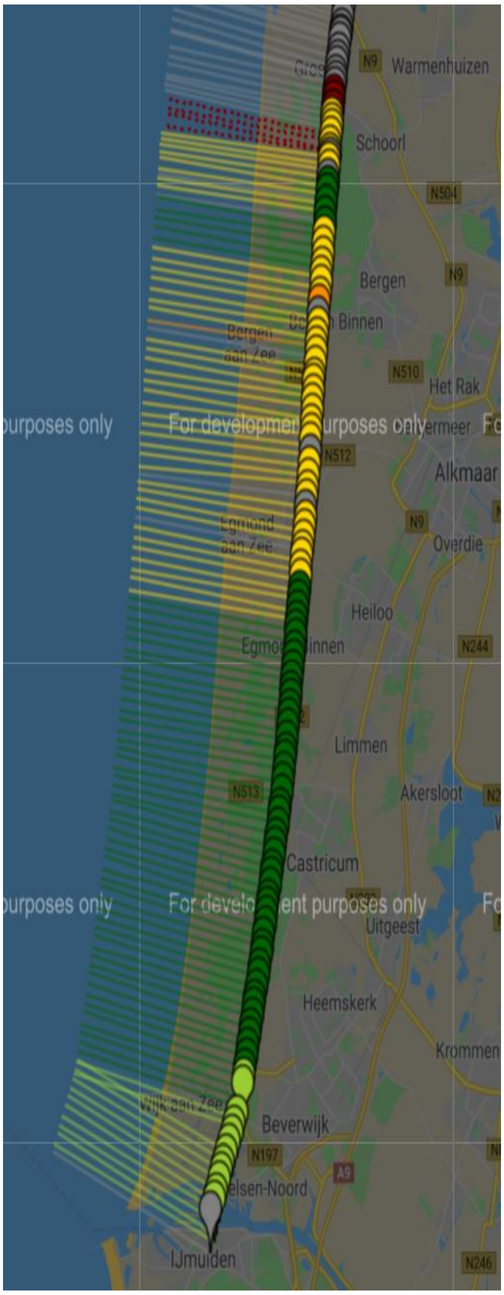


Figure IV.18: Graphical overview of the clusters in the southern part of the management area of HHNK. Solid line representing the main-clusters and the dashed line representing the sub-clusters.

[6] https://pgrwdhgzgnumatrk6qvssw-on.drv.tw/hkv/Clustering_Noorderkwartier.html

IV.4 Main-clusters

The first section of clusters, is dedicated to the main-clusters. These are clusters that contain the biggest part of the data set, containing more than 10 transects. The main clusters identify the bigger picture of processes and changes along the coastline of the management area of HHNK, where the sub-clusters can be seen as smaller processes. In total there are five main clusters isolated on the clustered U-matrix, as seen in figure IV.16.

Main-cluster 1 consists of 22 transects, located north of the HBD. These transects are bundled at one location. It represents the Callantsoog area which is represented by beach pavilions and recreational beaches. Main-cluster 2A consist out of 51 transects. The largest amount of these transects are located south of the HBD. There is also a small bundle of the same cluster, which is located north of the HBD. With south of the HBD a small and a large bundle. The northern location is located at Keeten with beach pavilions and recreational beaches. In the south the area represents both Egmond aan Zee and Bergen aan Zee. Main-cluster 3 contains 12 transects all bundled together. This cluster is represented by a large recreational area and is located at the most southern part of the HHNK coastline. Main-cluster 4A contains the largest amount of transects, with 77 transects in total. Most transects are located south of the HBD and a decent amount is located north of the HBD. Among all the clusters, main-cluster 4A is most spread among multiple locations along the HHNK coastline. Finally, main-cluster 5A consists of 19 transects. These transects are bundled together in the most north part of the HHNK coastline, representing recreational areas.

The next paragraphs will focus and explain the characteristics of each main-cluster. A full breakdown of the values for each cluster can be seen in appendix E. The average value as seen in the appendix are used. The main-clusters are explained with the use of the average values in appendix E and done in comparison to the other clusters. Highlighting the characteristics of each individual cluster. Table 7 represents a highlighted transect with the average values for each specific cluster.

Table 7: Partial representation of the average values in the main-clusters.

Cluster		**1**	**2A**	**3**	**4A**	**5A**
Amount of transects in cluster	**Unity**	**#22**	**#51**	**#12**	**#77**	**#19**
ActiveProfile_W	[m]	1016	1130	1000	863	689
Foreshore_W	[m]	824	918	768	719	433
Nourishmentforeshoreb	[-]	73.3	34.7	0.0	8.2	21.9
Nourishmentforeshoreb	[m^3]	2272	1074	0.0	255	677
docy	[NAP]	-8.9	-8.6	-5.4	-6.9	-11.9
foreshorem	[m^3/year]	8.55	5.33	4.05	1.63	5.01
Dune_W	[m]	192	213	232	144	257
DF_x2m	[m/year]	0.55	1.06	2.58	0.12	1.51
Bma_y	[NAP]	-0.6	0.0	1.4	1.1	-1.6
dunesm	[m^3/year]	2.99	3.69	4.53	1.51	5.56
beachwidthvar2b	[m]	77	97	149	85	78

IV.4.1 Main-cluster 1

Main-cluster 1 contains the highest dunes [19.2 m], which have been relatively stable [0.0005 m/year]. The crest of the dune has a slight seaward movement [0.54 m/year]. This seaward movement is also seen in the landward 6 meter location [0.77 m/year], the boundary between marine/aeolian transport [1.10 m/year] and a very large seaward foreshore movement [9.70 m/year]. While the dune foot [0.55 m/year] moves seawards, it is smaller compared to the other cluster. The derived height of the dune foot [0.03 m/year] largely increases. The intertidal width [65 m], beach width [77 m] and dune width [76 m] are in the same range. This is also seen in the yearly change of the intertidal width [0.23 m/year], beach width [0.83 m/year] and dune width [0.21 m/year]. The beach gradient [-0.037] is on the steep side, while both the dune gradient [-0.208] and foreshore gradient [-0.012] are less steep. The beach [0.0033] is rapidly flattening and the dunes [0.001] are relatively stable in their gradient.

Main-cluster 1 contains the largest foreshore nourishment [2272 m^3] and a decent amount of dune nourishment [782 m^3]. This is seen in the increasing foreshore volume [8.55 $m^3/year$] and dune volume [2.99 $m^3/year$]. While the active dune with [192 m] is small, the active foreshore width is large [824 m]. The depth of closure [-8.9 m] is located deeper than the average depth. At last, the Bma_y [-0.9 m] is below NAP. The highlighted transect 7001381 for main-cluster 1 can be found in figure IV.19.

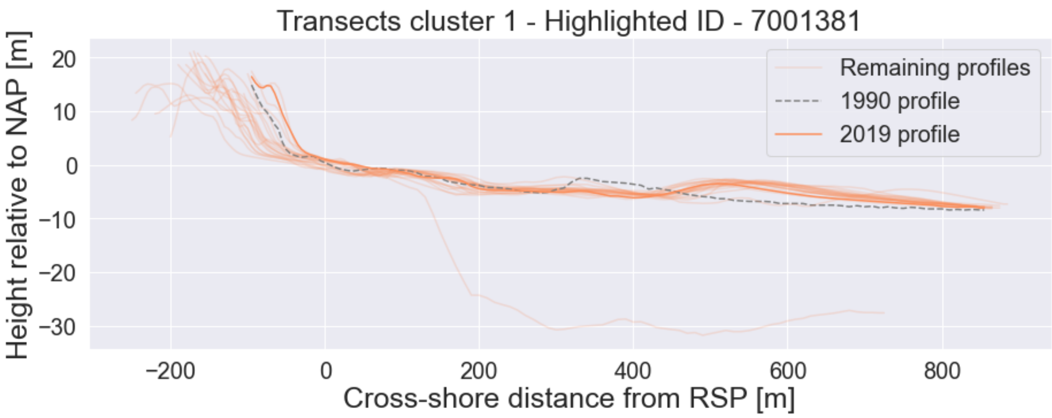


Figure IV.19: Main-cluster 1: All profiles and highlighted profile [7001381].

IV.4.2 Main-cluster 2A

Main-cluster 2A contains both high dunes [18.8 m] and short dunes [13.6 m]. The height of these dunes [0.024 m/year - 0.121 m/year] has been rapidly changing over the years, with a variety in seaward movements [-0.15 m/year - 0.50 m/year]. Both the landward 6 meter [0.95 m/year], the boundary between marine/aeolian transport [1.56 m/year] and seaward location [3.70 m/year], have a slightly higher seaward movement than average. The dune foot [1.06 m/year] also has an average seaward movement. Both the beach width [97 m] and intertidal width [82 m] are higher than average, The dune width [55 m] on the other hand, is on the smaller side. The dune gradient [-0.209] is average and both the beach gradient [-0.0037] and foreshore gradient [-0.024] are very steep.

The foreshore [2272 m^3] has received large nourishment, combined with an average dune [474 m^3] nourishment. Both the dune volume [3.69 $m^3/year$] and the foreshore volume [5.33 $m^3/year$] have a slightly larger increase than average. This is also reflected in the second-largest active profile width [1130 m] and a deep depth of closure [-8.6 m]. At last, the Bma_y [0.0 m] is around NAP. The highlighted transect 7003925 for main-cluster 2A can be found in figure IV.20.

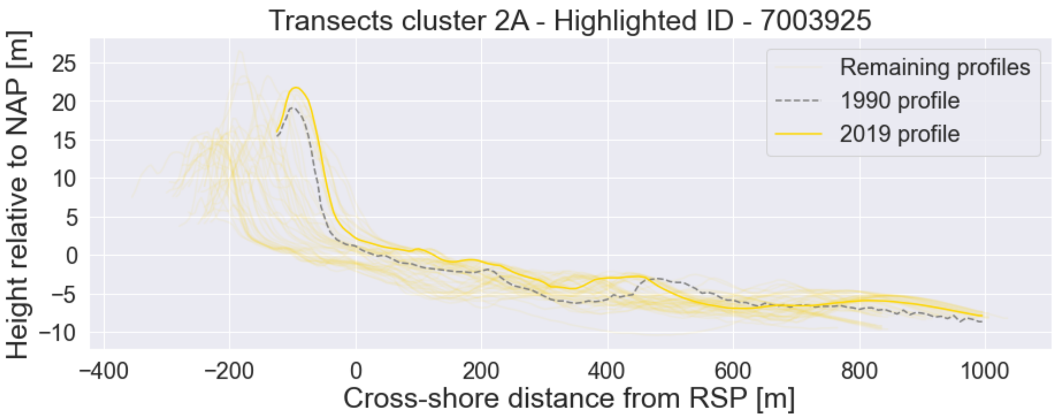


Figure IV.20: Main-cluster 2A: All profiles and highlighted profile [7003925].

IV.4.3 *Main-cluster 3*

Main-cluster 3 contains the highest variety in dunes. With high dunes [9.4 m] and short dunes [20.9 m]. On average, the height of the dunes [-0.036 m/year] decreases, with a very large seaward movement [2.27 m/year]. Besides the dune crest, the landward 6 meter position [2.45 m/year], the boundary between marine/aeolian transport [2.14 m/year] and seaward location [3.95 m/year] have a large seaward movement. With these large movements, the dune foot also has a very large seaward movement [3.95 m/year]. This cluster contains the largest intertidal width [107 m] and beach width [149 m]. All the gradients in this cluster are flat compared to the other clusters. Such as the beach gradient [-0.017] and foreshore gradient [-0.010]. The beach gradients is stable over time.

This clusters contains zero nourishment for both the foreshore [0.0 m^3] and dunes [0.0 m^3]. While there is no nourishment, there is a large increase in the foreshore volume [4.05 $m^3/year$] and dune volume [4.53 $m^3/year$]. The depth of closer [-5.4 m] is the most shallow. The active profile width [1000 m] is around the average of all clusters. At last, the Bma_y [1.4 m] is above NAP. The highlighted transect 7005450 for main-cluster 3 can be found in figure IV.21.

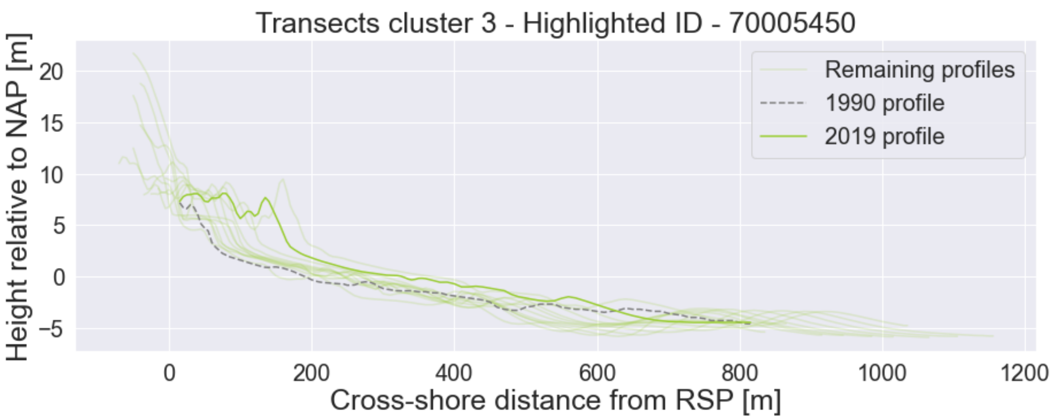


Figure IV.21: Main-cluster 3: All profiles and highlighted profile [7005450].

IV.4.4 Main-cluster 4A

Main-cluster 4A contains a large variety among the dunes, from small dunes [13.1 m] to high dunes [20.8 m]. This is also reflected in the large variety in dune crest movement [-0.03 - 0.86 m/year]. The height of the dunes [0.060 m/year] are increasing more than average. The position of the landward 6 meter location [0.23 m/year] and the boundary between marine/aeolian transport [0.38 m/year], have a very small seaward movement. The dune foot [0.12 m/year] shows negligible movement. The seaward foreshore location [1.95 m/year] movement is around the lower part of the spectrum. Both the intertidal width [77 m] and beach width [85] are average. While both the beach width [0.45 m/year] and intertidal width [0.45 m/year] increases by a small amount, the dune width [-0.13 m/year] shows to retreat. The dune front is steeper [-0.315] compared to most other clusters. The beach gradient [-0.030] is average and the foreshore gradient [-0.013] has a mild steepness. The beach gradient is getting more flatter [0.0013] and the dunes [-0.0014] more steeper.

This cluster contains one of the lowest amounts of foreshore nourishment [255 m^3] and dune nourishments [134 m^3]. This is reflected in the low increase in dune volume [1.51 $m^3/year$] and foreshore volume [1.63 $m^3/year$]. The active profile width [863 m] is among the shortest with a shallow depth of closure [-6.9 m]. At last, the Bma_y [1.1 m] is above NAP. The highlighted transect 7004750 for main-cluster 4A can be found in figure IV.22.

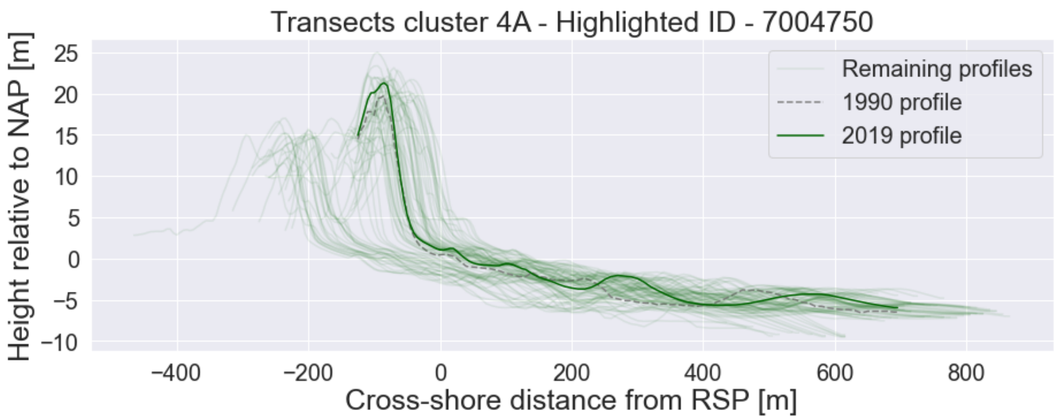


Figure IV.22: Main-cluster 4A: All profiles and highlighted profile [7004750].

IV.4.5 Main-cluster 5A

Main-cluster 5A contains a large variety from small dunes [14.8 m] to high dunes [17.6 m]. Both the change in height of the dunes [0.003 m/year - 0.085 m/year] and seaward movement [-0.01 m/year - 1.79 m/year] have a large variety. The intertidal width [53 m] and beach width [78 m] are among the smallest. The intertidal width [1.03 m/year] has a stronger change compared to the small change in beach width [0.42 m/year]. With the landward 6 meter [1.50 m/year], the boundary between marine/aeolian transport [1.73 m/year] and the dune foot [1.51 m/year] having a slightly larger than average seaward movement. The foreshore gradient [-0.020] and beach gradient [-0.037] are among the steepest. The beach gradient [-0.021] is getting steeper over time. Variables derived along the dune such as the dune foot [1.51 m/year], have a strong seaward movement.

This cluster contains the largest dune nourishment [1731 m^3] and a fair amount of foreshore nourishment [677 m^3]. Both the foreshore volume [5.01 $m^3/year$] and dune volume [5.56 $m^3/year$] have a large increase. The active profile width [689 m] is among the smallest. The depth of closer [-11.9 m] is located at the second deepest height. At last, the Bma_y [-1.6 m] is below NAP. The highlighted transect 7000409 for main-cluster 5A can be found in figure IV.23.

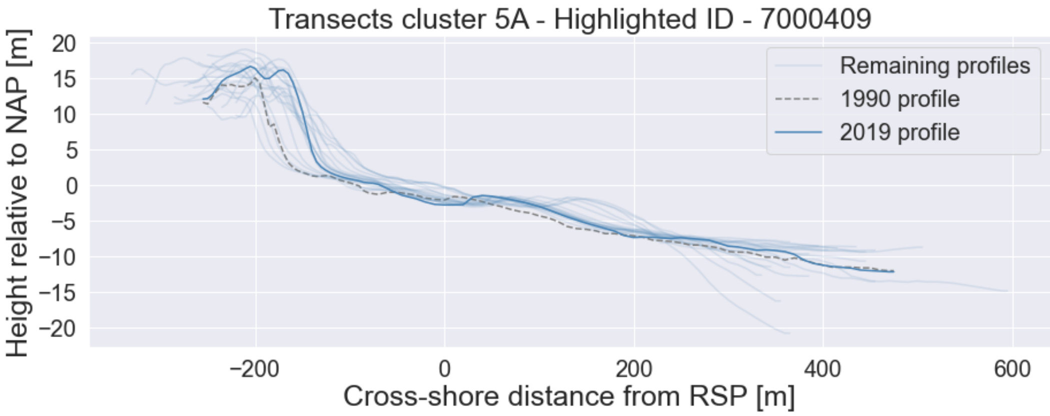


Figure IV.23: Main-cluster 5A: All profiles and highlighted profile [7000409].

IV.5 Sub-clusters

The second section of clusters, is dedicated to the "sub clusters". These are clusters along the HHNK coastline which stand out from the data analysis, but contain less than 10 transects. The sub-clusters are seen as anomalies in the clusters. Three out of the four clusters are classified as the same clusters by the K-means method, but were separated using the SOM. The annotation for these clusters contains a "B". This paragraph focuses on dissecting each of the sub-clusters. The dissection contains a visual representation of a highlighted transect in that particular cluster. Part of the data for the highlighted transect, can be seen in table 8. The table contains information about both the dominant variables and input variables, as spoken in chapter IV.6. The full table containing all input parameters and can be seen in Appendix E.

A quick overview of the clusters in figure IV.15 and figure IV.16, show that three out of the four sub-clusters are located in the northern part of the management area of HHNK. To further research each sub-cluster, the effect of both the clustering method is examined, as seen in chapter IV.2.2. Increasing the number of centroids from four to six, doesn't change the boundaries, as seen in the results with four clusters. It can also be seen that cluster 0 remains stable with an increasing number of centroids and doesn't have any other corresponding connection with other main clusters. The sub-clusters are strongly identified using the SOM, as seen in figure IV.8, compared with the clustered U-matrix in figure IV.15.

Sub-cluster 0 contains four transects and is located close to the HBD. While the final data set has the HBD removed, it relied on the reduced data as seen in chapter III.2.2. The key concept here was based on measurement quality before the year 2001. Meaning the remnants, are the transects that are included in the nourishment of the HBD and have measurements up till 1990. Sub-cluster 2B is located at the most northern location in the management area of HHNK. This cluster contains five transects. Sub-cluster 4B is the second-highest cluster in the north, containing eight transects. The last sub-cluster 5B, is located at the most southern location in the northern part of the management area of HHNK. It can be noted that each of the sub-clusters is located in the north.

In the following paragraphs, the average values of the clusters are explained and compared to all other clusters.

Table 8: Partial representation of the average values in the sub-clusters.

Cluster		**0**	**2B**	**4B**	**5B**
Amount of transects in cluster	**Unity**	**#5**	**#5**	**#8**	**#3**
ActiveProfile_W	[m]	1100	1554	869	643
Beach_W	[m]	387	1300	648	508
Nourishmentforeshoreb	[-]	12.7	62.7	16.1	12.7
Nourishmentforeshoreb	[m^3]	393	1943	498	393
docy	[NAP]	-10.1	-19.8	-8.3	-6.9
foreshorem	[m^3/year]	6.64	-6.20	3.85	1.98
Dune_W	[m]	713	254	222	135
DF_x2m	[m/year]	0.34	1.55	1.70	-0.64
Bma_y	[NAP]	-5.6	-2.8	0.1	0.4
dunesm	[m^3/year]	5.13	3.56	5.18	0.35
beachwidthvar2b	[m]	102	70	88	66

IV.5.1 Sub-cluster 0

The dune nourishment [1099 m^3] in sub-cluster 0 contains mean nourishment which is larger than both the 25th percentile [254 m^3] and 75th percentile [396 m^3]. This originates from a certain transect that contains nourishment information from the HBD. Both the nourishment and effect of the HBD, is seen in the increase in dune volume [5.13 $m^3/year$], which is the largest compared to all clusters.

The dune crest has a large variety between both high dunes [19 m] and low dunes [14 m]. The height increase of the crest [0.013 m/year] is among the smallest. This cluster, is the only cluster with a landward moving dune crest [-0.32 m/year]. The intertidal width [77 m] and its changes [0.38 m/year] are average. The beach width [102 m] is the second largest and has the largest increase in width [2.67 m/year]. The boundary between marine/aeolian transport [2 m/year] and seaward foreshore location [9.32 m/year], has a high seaward movement. The landward 6 meter location [0.29 m/year] has a very small seaward movement. The isolated height of the dune foot [0.03 m/year] moves upwards, affecting the horizontal isolation of the dune foot. The dune foot [0.34 m/year] has a small seaward movement. This cluster contains one of the smallest dune width [50 m], which largely increases over time [0.87 m/year]. Both the beach gradient [-0.029] and the foreshore gradient [-0.0015] are not very different from most clusters. The beach gradient however, becomes more flatter [0.00052] than any other cluster. The most interesting value, is the dune gradient [-0.363]. This gradient is the steepest and rapidly changes [0.0067] to a flatter gradient.

At last, the changes of the dune volume [5.13 $m^3/year$] and the foreshore volume [6.64 $m^3/year$], are among the highest. The active beach width [387 m] is among the smallest, with the cluster containing the largest dune width [713 m] is the largest. Resulting in a large total active profile width [1100 m]. The depth of closure [-10.1 m] is at a very deep height. At last, the Bma_y [-5.6 m] is located at the lowest height below NAP. The highlighted transect 7002782 for sub-cluster 0 can be found in figure IV.24.

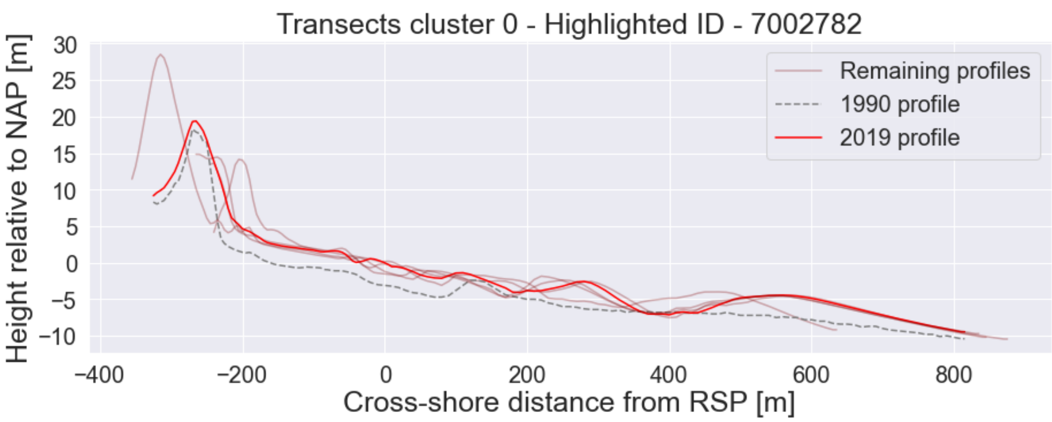


Figure IV.24: Sub-cluster 0: All profiles and highlighted profile [7002782].

IV.5.2 Sub-cluster 2B

Both the height [16.9 m] and changes [0.046 m/year] of the dune crest in sub-cluster 2B are average. The seaward movement [0.12 m/year] is small. Both the landward 6 meter [1.01 m/year] and the boundary between marine/aeolian transport [1.28 m/year] have the same order of magnitude. While the foreshore location [2.09 m/year] moves seawards much faster, the locations are average. The dune foot [1.55 m/year] also moves in the same order and doesn't change height. The intertidal width [46 m] along this cluster is the smallest. The beach width [70 m] is also very small. Besides the beach width being small, it is also the only cluster that shows a retreat of the beach width [-0.10 m/year]. Both the intertidal width [0.60 m/year] and dune width [0.61 m/year] are growing. Besides the decreasing beach width, the beach gradient [-0.037] and the foreshore gradient [-0.024] are very steep. The beach gradient is gaining more steepness over time.

These changes in certain variables can be explained, due to this being the only cluster that has a very large decrease in foreshore volume [-6.20 $m^3/year$]. The dune volume [3.56 $m^3/year$] however, is steadily increasing. With a large amount of foreshore nourishment [1943 m^3] and a very large active foreshore width [1300 m], the depth of closure [-19.8 m] lies very deep. The active dune width [254 m] is in the same order of magnitude as most clusters. At last, the Bma_y [-2.8 m] is below NAP. The highlighted transect 7000230 for sub-cluster 2B can be found in figure IV.25.

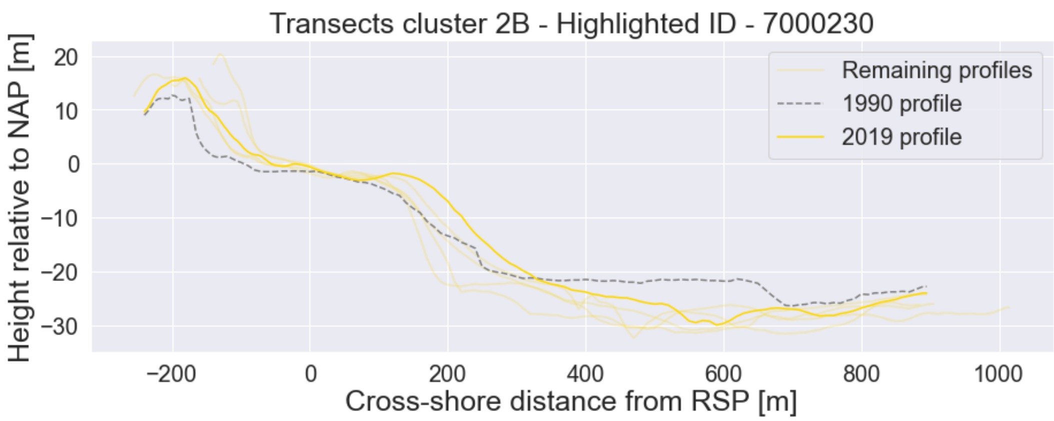


Figure IV.25: Sub-cluster 2B: All profiles and highlighted profile [7000230].

IV.5.3 Sub-cluster 4B

Sub-cluster 4B contains the smallest dune crest [14.9 m] and is largely increasing [0.127 m/year], with a steady seaward movement [0.41 m/year]. Both the landward 6 meter [1.43 m/year] and the boundary between marine/aeolian transport [1.51 m/year] has the same order of magnitude moving seawards. The seaward foreshore location [2.30 m/year] moves almost twice as fast seawards. The height of the dune foot stays the same, while the horizontal movement is strongly seawards [1.70 m/year]. The dune width [63 m] is relatively small compared to all other clusters, but has been increasing very fast [1.01 m/year]. Both the beach width [88 m] and its changes [0.36 m/year] are average. The beach gradient steepens in time [-0.04].

With hardly any dune nourishment [60 m^3] and foreshore nourishment [498 m^3], there is a large increase in dune volume [5.18 $m^3/year$] and a decent increase in foreshore volume [3.85 $m^3/year$]. The total active profile width [869 m] is average with a slightly deeper depth of closure [-8.3 m]. At last, the Bma_y [0.1 m] is around NAP. The highlighted transect 7000869 for sub-cluster 4B can be found in figure IV.26.

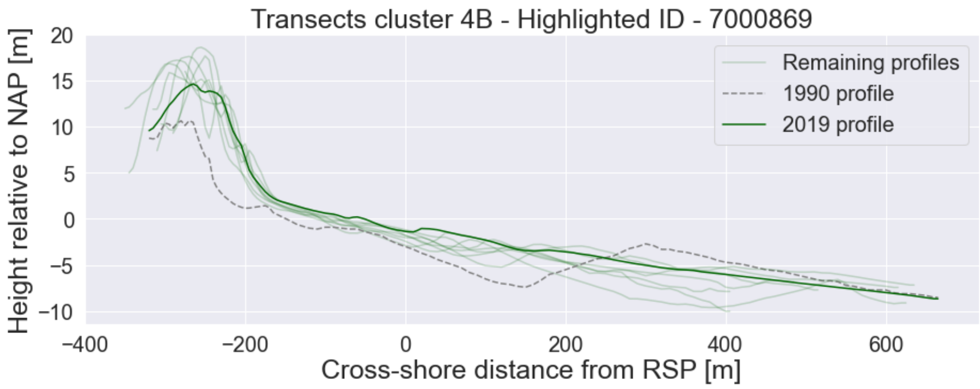


Figure IV.26: Sub-cluster 4B: All profiles and highlighted profile [7000869].

IV.5.4 Sub-cluster 5B

The height of the dune crest [15.8 m] in sub-cluster 5B, is lower than average and increases steadily [0.025 m/year]. The location of the dune crest [1.62 m/year] moves seawards. A unique characteristic compared to all other clusters, is that the landward 6 meter landward position [-0.35 m/year] and the boundary between marine/aeolian transport [-0.35] are retreating. In contrary, the seaward foreshore location [2.66 m/year] moves seawards. The dune foot [-0.64 m/year] is retreating as well. Resulting in a decrease in dune width [-2.04 m/year]. This retreat causes the dune steepness to increase strongly [-0.0087]. The beach width [66 m] is one of the smallest, but is increasing [0.76 m/year] steadily. Compared to the other clusters, the beach gradient [-0.038] and the foreshore gradient [-0.017] are average.

Both the dune nourishment [353 m^3] and the foreshore nourishment [393 m^3] are very small compared to all other clusters. The foreshore volume [1.98 $m^3/year$] and dune volume [0.35 $m^3/year$] have a very small increase. This cluster also contains the smallest active profile width [643 m] and a shallow depth of closure [-6.9 m]. At last, the Bma_y [0.4 m] is around NAP. The highlighted transect 7002782 for sub-cluster 5B can be found in figure IV.27.

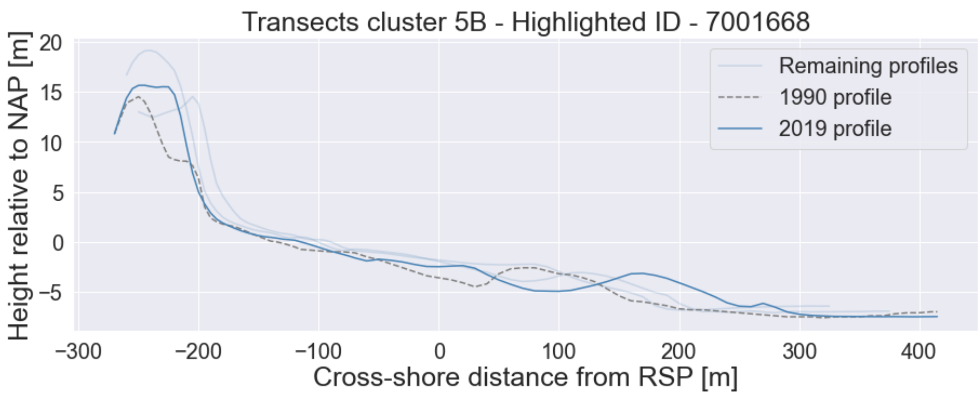


Figure IV.27: Sub-cluster 5B: All profiles and high profile [7001668].

IV.6 Dominant variables

IV.6.1 Active Foreshore Width

Nine clusters are identified in the clustered U-matrix in the final phase. After combining the SOM and the K-means algorithms, an extra step was taken in this research. This step plots the contours of the clustered U-matrix back onto each individual two-dimensional component plane. This results in a clear and fast visualisation of the dominant variables which identify the clusters. At the same time, each individual cluster can be inspected on its own unique characteristics relative to all component planes of each variable. The first degree of dominant variables, are the total active profile width ('ActiveProfile_W') and active foreshore width ('Foreshore_W'), which can be seen in figure IV.28. The total active profile width is dependent on two components, the active dune width and active foreshore width.

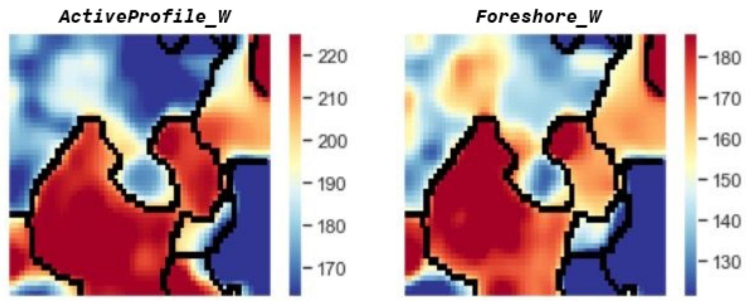


Figure IV.28: First degree of dominant variables for the final data set. Represented by the total active profile width (ActiveProfile_W) and the active foreshore width (Foreshore_W).

As the clustering is dependent on high-dimensional data, there is more depth in the reason for the clustering than a single variable. The second level of dominant variables which follow the contours of the clustered U-matrix, but have slightly more deviation, are the amount of foreshore nourishment ('Nourishmentforeshoreb'), depth of closure y-coordinate ('docy') and the change of total foreshore volume ('foreshorem'). Besides comparing the individual variables with the clustered U-matrix, the individual component planes can also be compared with each other to identify correlations for specific clusters. The full set of variables combined with the contours of the clustered u-matrix, can be found in appendix C.

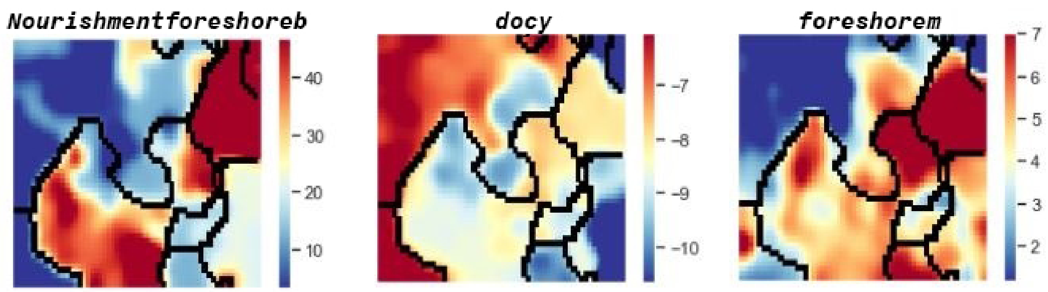


Figure IV.29: Second degree of dominant variables for the final data set. Represented by the foreshore nourishment (Nourishmentforeshoreb), y-coordinate of the depth of closure (docy) and change in foreshore volume (foreshorem).

Having isolated the dominant variables, the variables are compared with each other. This includes the foreshore nourishment ('Nourishmentforeshoreb') and the dunes ('Nourishmentdunesb'). The nourishments are compared with the changes of volumes in the foreshore ('foreshorem') and the change of volume in the dunes ('dunesm'). These four variables are visualised in figure IV.30 and plotted with their values on the left. The largest dominant variable such as the active foreshore width and ('Foreshore_W'), is seen as a red line in figure IV.30.

The foreshore nourishment and the active foreshore width, show to follow the same pattern for most locations from north to south. The locations where this pattern isn't seen, such as cluster 0, are also identified in figure IV.29. In chapter II.5.3, an experimental setup was used to identify the combined effects of long-term changes such as sea level rise and nourishment. The results in this research correspond to the same results as the experiment for most clusters. With nourishment resulting in a shifting of the depth of closure and therefore resulting in a larger total active width.

To further investigate this effect, two transects are compared. The first transect 7003425, has a large amount of nourishment and a large active beach width, as seen in figure IV.30. The second transect 7004500, has a very small amount of nourishment and a small active beach width, as seen in figure IV.30. In both examples the standard deviations are plotted, cumulative for each year from 1995 to 2019. This indicates if the depth of closure would be derived at a different location if more or fewer years are used. While using this method to compare the effect of nourishment, having more years equals a more qualitative standard deviation Hinton [2000].

For the first transect 7003425, the results are seen in figure IV.31. The derivation of the depth of closures moves seaward, with a total amount of 237 meters over the years. Comparing this to the standard deviation of the second transect 7004500, the depth of closure stays at the same place over the years, as seen in figure IV.32. Comparing both the results of this research and the experiment done by Aktinson and Baldock [2020], a correlation between the depth of closure and therefore the active profile width due to nourishment is found.

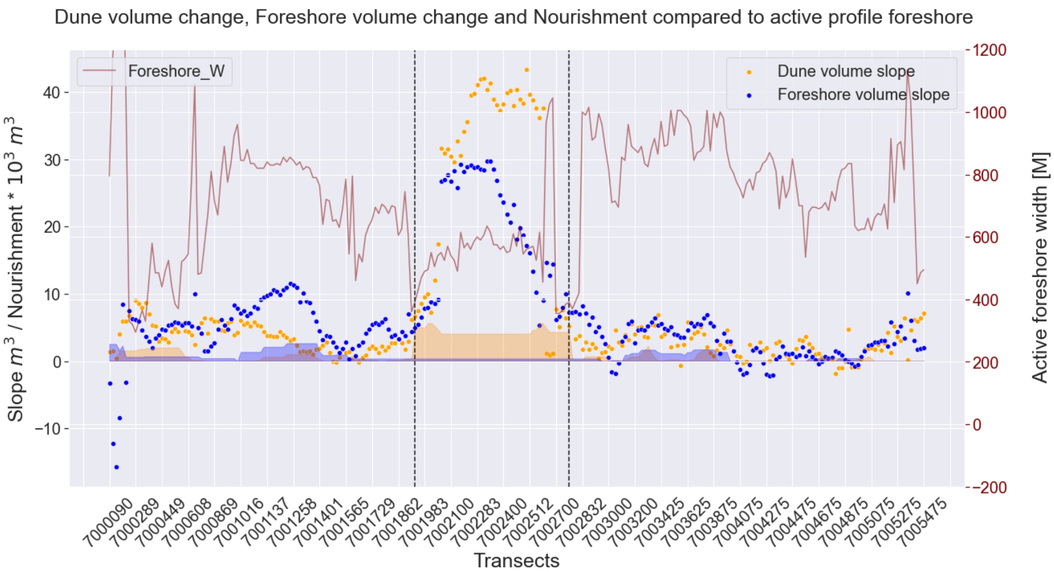


Figure IV.30: Representation of variables from the north to the south of the management area of HHNK. Red line representing the active foreshore width, the blue dot representing the change in foreshore volume, the orange dot representing the change in dune volume, the blue shaded area representing the foreshore nourishment and the orange shaded area representing the dune nourishment.

Besides the depth of closure, the standard deviation method is also used to derive the landward boundary and the Bma_y. The landward boundary, located at the crossing between the STD threshold left of the standard deviation peak in the dunes, show almost no differences between the first couple of years and the years after. The landward boundary is located at around -230 meters for figure IV.31 and -30 meters for figure IV.32, relative to the RSP. The Bma_y, which is the minimum between the standard deviation peak in the dunes and the standard deviation peak in the foreshore, is slightly harder to find. But in both cases almost no differences between the first years and the years after are found.

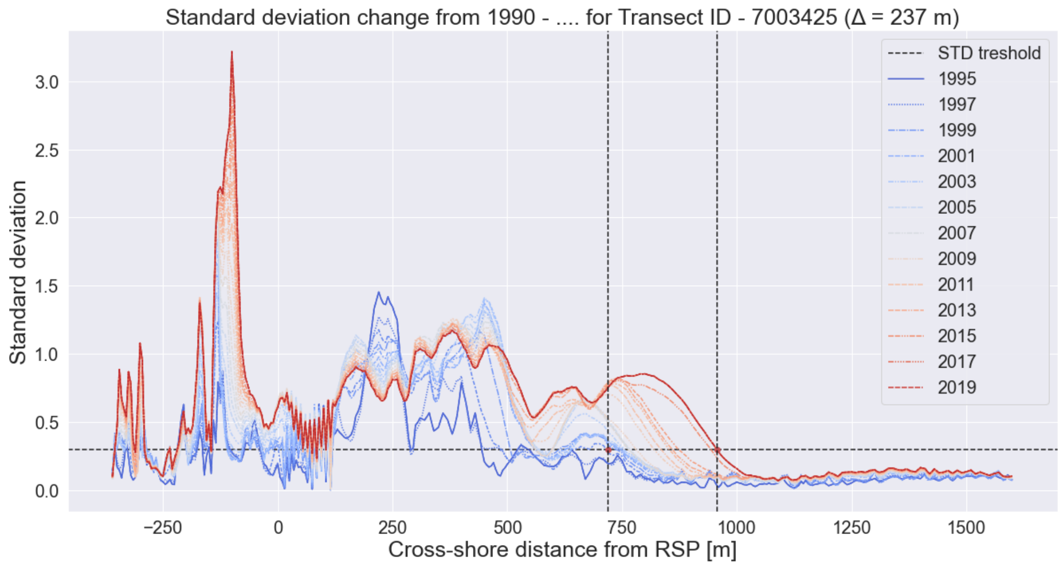


Figure IV.31: Standard deviation for different cumulative years for transect 7003425. With a shifting depth of closure of 237 meters seawards.

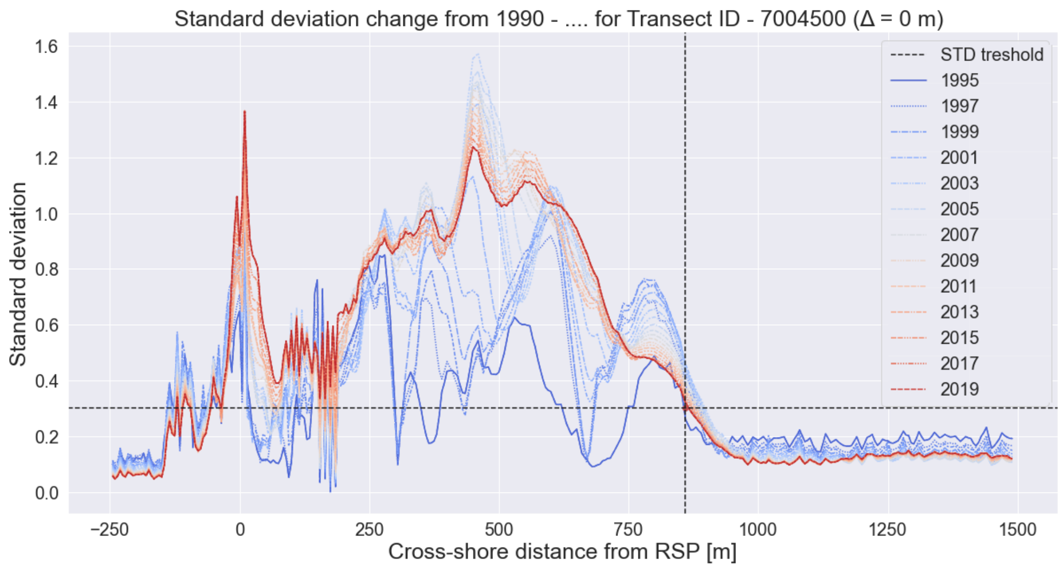


Figure IV.32: Standard deviation for different cumulative years for transect 7004500. With no shifting depth of closure.

IV.6.2 *Correlation in dune variables*

To identify the trend in horizontal change of the depth of closure and nourishment in the identified clusters, the derivation for the depth of closure is slightly modified. This is done due, because an increase in cumulative measurements results in a more qualitative standard deviation profile. With fewer measurements, the standard deviation profile is less representative of the changes in the profile. This results in a paradox where fewer measurements give valuable information but at the same time are qualitative less representative of the standard deviation profile. Therefore the algorithm is adjusted and the threshold is set higher to 0.5 to remove smaller anomalies. Due to each cluster having different processes which are dominant, the standard deviation profile can also show general differences. Figure IV.33 represents the total change in distance of the depth of closure (e.g. figure IV.31) in regard to the total amount of nourishment for each individual transect in each cluster.

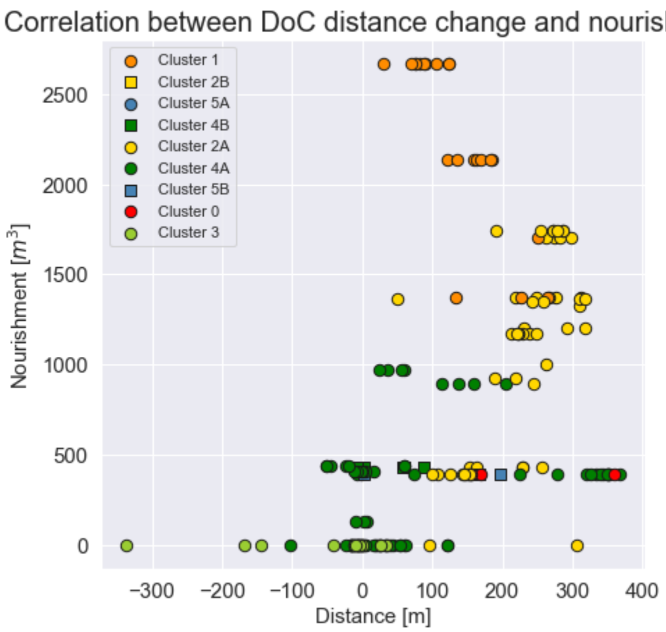


Figure IV.33: Correlation graph with the amount of nourishment and the shifting depth of closure for each transect in each identified cluster.

The correlations in figure IV.33 can be classified by their clusters and is found in table 9. Indicating that for most clusters, the correlations can't be found, especially in the sub-clusters. This is due to no nourishment being present or the small number of data points. Both cluster 1 and cluster 2A are associated with a high nourishment, while cluster 4A is associated with almost zero nourishment. This results in biases in the derivation of the correlations. With the identified characteristics of each cluster, the correlation gives valuable information.

Table 9: Correlation (Pearson, Spearman, Kendall) for all identified clusters.

Correlation/Clusters	**0**	**1**	**2A**	**2B**	**3**	**4A**	**4B**	**5A**	**5B**
Pearson	-	-0.82	0.62	-	-	0.34	-	-	-
Spearman	-	-0.84	0.61	-	-	0.26	-	-	-
Kendall	-	-0.70	0.48	-	-	0.18	-	-	-

The strongest correlation is found in cluster 2A. This cluster represents a large part of the management area of HHNK with 51 transects. As this is not the only cluster that is nourished, it is a particular cluster inside all nourished locations. Where cluster 1 is also classified as a nourished transect, it represents a negative correlation, with a high probability that this originates from the ebb-tidal delta.

Removing small anomalies from the data set, results in a very strong Pearson correlation of 0.732, as seen in figure IV.34. Indicating there is a high correlation between the nourishment and the horizontal movement of the depth of closure, particular in cluster 2A.

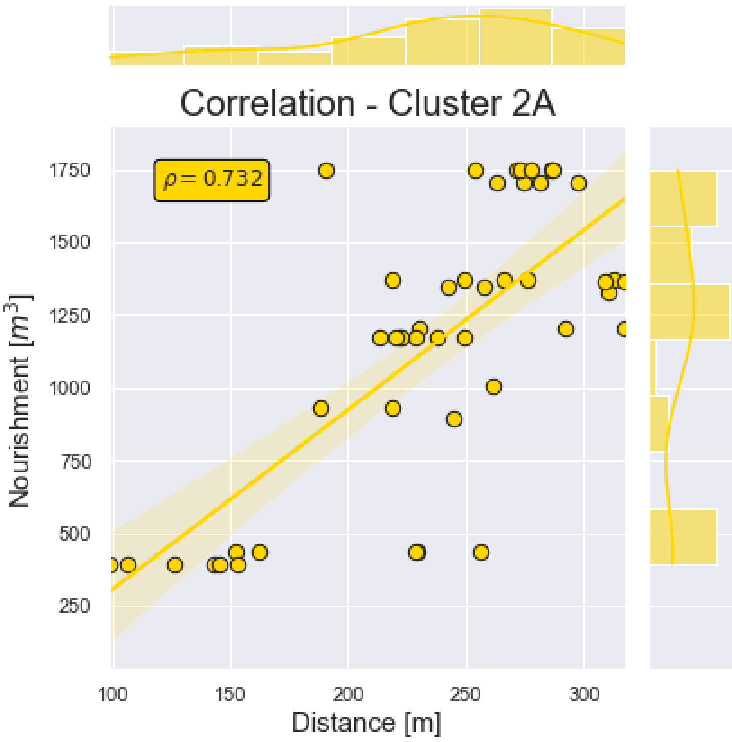


Figure IV.34: Calculated Pearson correlation for the data isolated for main-cluster 2A.

IV.6.3 Active Dune Width

In the previous paragraph, the active foreshore width is seen as the main dominant variable for the clustering. This paragraph will focus on the correlation between variables that do not dominate the clustering. The total active profile width consists of both the active dune width and the active foreshore width. The active dune width is defined as the width between the landward stable point and the Bma_y, as explained in chapter III.5.1.

Figure IV.35 shows the correlation between variables from the input space on the component plane. A correlation is seen between the active dune width, horizontal change of the dune foot, Bma_y and the change in dune volume. Besides these variables, the Bma_x and landward 6 meters also correlate with these variables. The clustered component planes of each individual variable can be seen in Appendix C.

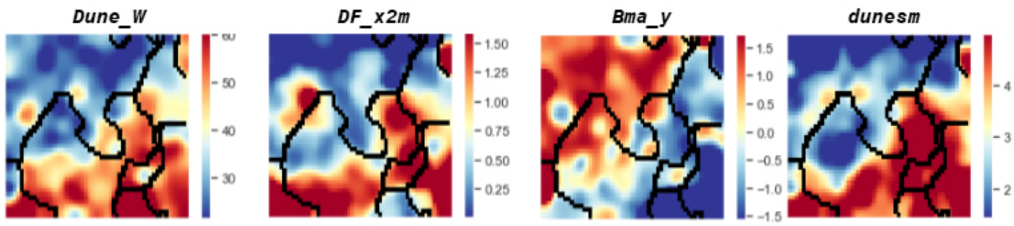


Figure IV.35: Clustered component planes for the active dune width (Dune_W), Dune foot horizontal change (DF_x2m), y-coordinate boundary marine/aeolian transport (Bma_y), dune volume change (dunesm).

The correlations show that for most clusters, a larger active dune width correlates with a larger dune volume increase. Though the derivation of the dune volume is based on the active dune width, the derivation of the horizontal seaward movement of the dune foot, isn't based on any correlated variable. At last, the Bma_y lies above NAP for the locations with a small change and lies below NAP for transects with a large change.

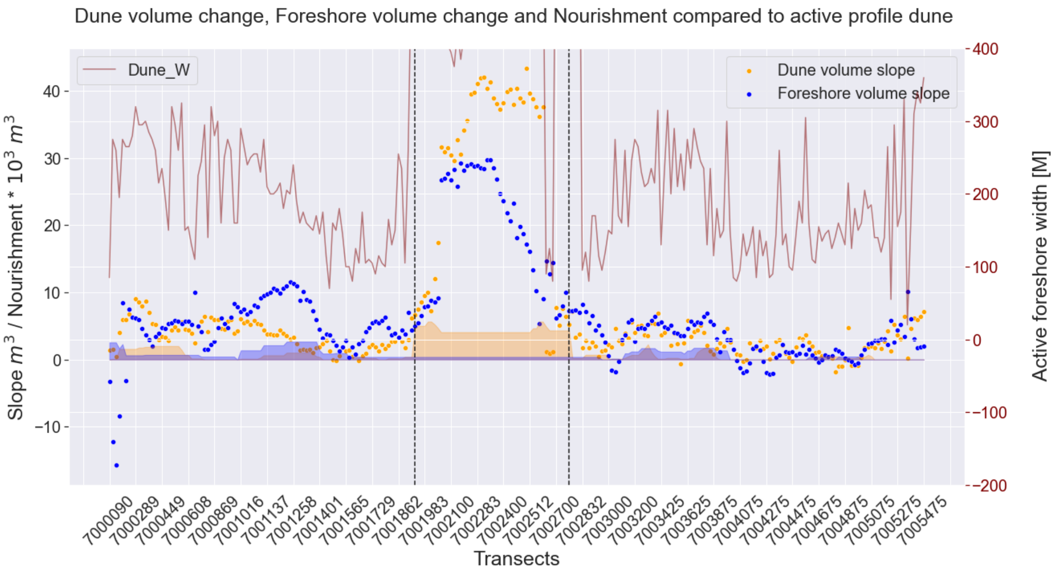


Figure IV.36: Representation of variables from the north to the south of the management area of HHNK. Red line representing the active dune width, the blue dot representing the change in foreshore volume, the orange dot representing the change in dune volume, the blue shaded area representing the foreshore nourishment and the orange shaded area representing the dune nourishment.

V. Discussion, Conclusion and Recommendations

This chapter starts with paragraph V.1, in which various parts of the research are discussed. Followed, by the conclusion of this research in paragraph V.2. This research ends in paragraph V.3, containing recommendations for HHNK on how to implement and expand further upon the results of this research, into a better understanding of the complex coastal zone.

V.1 Discussion

V.1.1 Data Collection

The most important aspect of any data analysis, is to understand the source of the data. The first part of the data collection, consists of the derived variables from the JarKus data set. Variable are derived based on an algorithm or method, each of them containing a certain bias. The definitions for variables, such as the beach width, can be derived as a fixed or variable location, both representing the beach width with different results. While the bias is minimized by taking a regression line and clustering the data, it needs to be taken into account when interpreting the results. The second part of the data collection, consists of the nourishment data. This data is provided by Rijkswaterstaat and contains a starting transect, ending transect and the total amount of volume nourished. The total volume of nourishment is evenly spread along all the transects. But it is practically impossible to nourish each transect with the same amount of volume. The last part of the data collection contains variables, that are derived with the standard deviation method. The standard deviation method has previously only been used to identify the depth of closure. Within this research, it is extended for locations, widths, gradients and volumes. For the most part it showed great potential for deriving multiple variables. But there are transects that have a hard time identifying the depth of closure. For example, the northern part of the management area of HHNK, showed to have more trouble with the derivation than the southern area. This was one of the reasons for removing certain transects.

V.1.2 Method

The K-means algorithm is one of the most used unlabelled clustering techniques, while the SOM is used on a much lower scale. This research combines both algorithms and reflect the results on the input space. This approach is used on a even much lower scale. Therefore, this research focuses also on the applications of the hybrid algorithm and its results with respect to different input data. As no guidelines are present for the combination of both algorithms, the guidelines were chosen based on the fundamentals of both algorithms, as both have the same goal but a different approach. The SOM can be seen as a constrained K-means. In the final data set, the elbow method showed to be insufficient. This is due to the effect that adding centroids in the K-means does not change a large part of the clusters. After interpreting the results, it is also seen that another approach could be. to look at the clusters in the spacial location of the management area of HHNK and take every boundary from north to south as an individual cluster. At last, reflecting the clustered U-matrix onto the component planes, revealed the dominant variables. The SOM gives great insight into the complex high dimensional data in a two-dimensional plane. Even though this gives more insight for human interpretation, translating the results remains complex when looking into the "why" aspect of the results. This is why machine learning is often referred to as a "Black Box" approach. While this applies to this research as well, the results gave great insight into part of the "why" question.

V.1.3 Results

The results of this research can be categorized in three sections; Clustering of the management area of HHNK, horizontal change of the depth of closure and correlation in dune characteristics.

Clusters

The management area of HHNK is classified in nine clusters, based on unlabelled data. While this clustering reveals interesting results, it is based on the intelligence of a machine and the input data. The clusters are based on their foreshore changes, which might not be in line with morphological understanding of the coast. As the foreshore changes are largely based on human interventions, the clusters give new perspective of the coast, based on the data analysis. The most southern cluster is most likely defined by the pier at IJmuiden. In the North, the ebb-tidal delta most likely defines the clusters. While these exact processes aren't researched, the clustering algorithm isolates part of the coastline changes due to these influences being present. The larger clusters show more range of variability of the results, but still contain their unique characteristics. Each of the these clusters could be made smaller, to identify more local changes.

Depth of Closure

The results show that the active foreshore width, foreshore nourishment, foreshore volume change and the depth of closure, are the dominant variables for the clustering. Upon further investigation, a strong correlation between the horizontal shifting of the depth of closure and nourishment, was found in nourished clusters. These results correspond to the experimental research Aktinson and Baldock [2020]. While the experimental setup wasn't intended on finding the depth of closure, the horizontal shifting effect was clearly seen. These results are important to the current policies and understanding of the complex coastal zone. The results affect two main components of the current policies. The first component is, that the depth of closure is used to determine the amount of nourishment needed at the coast and is an offshore limit for both numerical models and fill designs Morang and Birkemeier [2005] Rijkswaterstaat [2020]. The second component, addresses the ecological effect of nourishment. The biodiversity on the bottom of the sea is distorted due to nourishment Rijkswaterstaat [2020]. The specific effect is complex and location-dependent. The results of this research can be used in future ecological research off the effect of nourishment. There has been previous research that indicates that the depth of closure is variable alongshore and its derivation is influenced by bar behavior Marsh et al. [1998]. With this research including more qualitative and quantitative data of the JarKus data set, it can now also be concluded that the depth of closure can also be variable in time.

Boundary between marine and aeolian transport

The boundary between marine and aeolian transport, indicates till what location the marine processes and aeolian processes are present De Vries et al. [2010]. A clear pattern is seen between the location of this boundary, relative to NAP and the variables along the dunes. The correlation is found between the active dune width and increase in dune volume, which originates from the same standard deviation method. Therefore the correlation and variables could be correlated due to the derivation method. However, it also correlates with the horizontal movement of the dune foot, which isn't derived with the standard deviation method, validating the correlation to some degree. The question that remains is, if the correlation is equal to the causation. In this research it's too early to make this statement, it does however provide a strong underlying argument for further research. Presuming that, a lower Bma_y results in a larger supply area for aeolian transport and therefore a larger dune growth and vica versa. As this research didn't take wave climate into account, further research in the origin of the Bma_y by means of standard deviation is advised.

V.2 Conclusion

V.2.1 Research questions

To answer the main research question, the following research questions have been answered:

- **How can long term changes be characterized along the coastline of the HHNK area, using advanced data analysis and machine learning techniques ?**

  The long-term changes of the management area of HHNK, derived from the JarKus and nourishment data, are bundled together in a high-dimensional unlabeled data set. Research on advanced data analysis techniques, concluded that clustering techniques, which are a branch of machine learning, are best fitted to understand this complex high-dimensional data set. The K-means algorithm is used in combination with the Self-Organizing Map. The algorithms are combined to form a hybrid algorithm that has the ability to plot the structures of the clustering, back onto each individual component plane of the input space variables. This results in identifying similar unique characteristics for each cluster of long-term changes along the coastline of HHNK.

- **Which locations along the coastline of the HHNK area, can be clustered in time and space with their associated characteristics ?**

  The results of the initial data sets show, that the clustered transects, are stable in time and vary in space. Using a regression line, the information in time is contained in a single value. In the final data set, the derivation of volumes is adjusted, as it shows to be the dominant variable in the initial data set. With these adjustments made from the initial data set, the final data set identifies nine clusters along the coastline of HHNK. Four of these clusters have a large number of transects and are denoted by main-clusters. The remaining clusters with a low number of transects are denoted by sub-clusters. The dominant variable, which identifies the similarities among individual clusters, is the active foreshore width. This is combined with lesser dominant variables such as the foreshore nourishment, depth of closure and changes in the volume of the foreshore. With one highlighted transect from all nine clusters to be seen in figure V.1. The unique characteristics of the main-clusters and sub-clusters are as followed:

  Main-Cluster 1: The highest and most stable dunes. With the largest amount of foreshore nourishment, the foreshore volume has the largest increase in volume. Decent amount of dune volume, with the dunes increasing in volume. The dunes are moving seaward combined with a flattening of the beach and the foreshore.

  Main-Cluster 2A: Large amount of dune and foreshore nourishment. Dunes have a large variety in both height and gradient. The beach width and intertidal width are large and the dune width is small. Both the beach gradient and foreshore gradients are very steep. The increase in volume is slightly larger, compared to all clusters.

  Main-Cluster 3: On average the dune crest is decreasing and has a large seaward movement. Largest widths and most flat gradients. With zero nourishment, both the foreshore volume and dune volume are increasing. The depth of closure is the most shallow of all clusters.

  Main-Cluster 4A: One of the lowest nourished clusters, with a very small increase in volume. Dune foot shows negligible movement, with very small seawards movement for most variables. With this cluster containing the most transects, it also shows the most variety in most variables.

Main-Cluster 5A: Contains the largest amount of dune nourishment and a decent amount of foreshore nourishment. The depth of closure is located at the second deepest height, with a very small active profile. The gradients are among the steepest, and are steepening over time. The widths in this clusters are among the smallest. The dune foot has a strong seaward movement.

Sub-Cluster 0: Contains remnants of the HBD nourishment which is removed in the final data set. Large beach width, with small and steep dunes that are flattening. Large active dune width and a small beach width. Retreating dune crest and small seawards movement of the dune foot, but large movement of the seaward foreshore location. Very large increase in foreshore volume and dune volume.

Sub-Cluster 2B: Small seaward movement of the dune crest and larger seaward movement of the dune foot. Largest beach nourishment and very high dune nourishment. Contains the deepest depth of closure and largest active profile. Very small beach and intertidal width, with a steep beach gradient. The only cluster with retreating beach width, steepening beach gradient and decrease in foreshore volume.

Sub-Cluster 4B: Average dune crest has the largest increase in height. The largest increase in intertidal width, dune width and dune foot. With almost no dune nourishment and foreshore nourishment, the increase in dune volume and foreshore volume are one of the largest.

Sub-Cluster 5B: Very small intertidal width and beach width, with the steepest beach gradient. Shallow depth of closure, and smallest active profile width. A small amount of dune and foreshore nourishment. The dune crest has a large seaward movement and is the only cluster with a retreating landward 6 meter location, the boundary between marine/aeolian transport and dune foot. The dune width is also decreasing and largely steepening. No increase in dune volume and a small increase in foreshore volume.

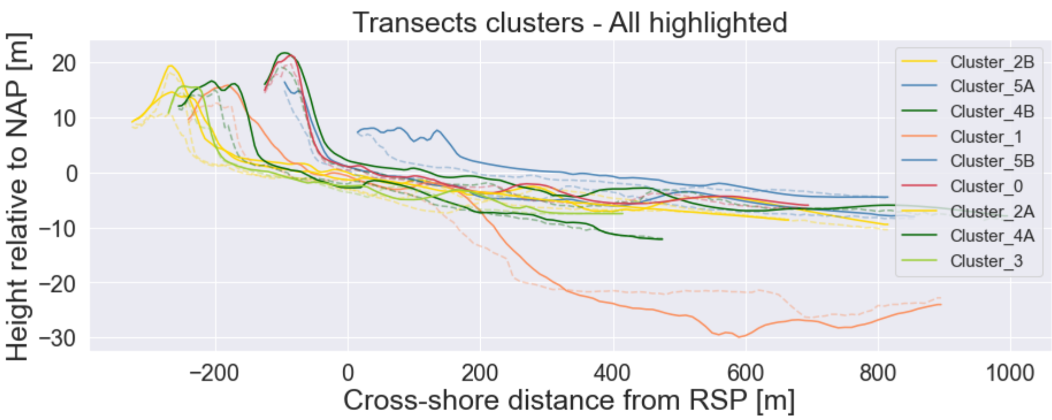


Figure V.1: Graphical representation of active profile for the highlighted transects of the nine clusters.

- **How do the long term changes along the HHNK area compare with physical understanding of the long term coastal processes ?**

  The identified nine clusters and their similarities, are the result of both long term natural processes and human interventions. Each cluster is represented by different changes in the coastal profile and therefore by their own long-term changes. Relating the unique characteristics in each cluster, to their own unique combination of natural processes and human interventions. The similarities between clusters, were found to be dominated by the active foreshore width, the foreshore nourishment, depth of closure and increase in foreshore volume. The standard deviation method explains this correlation, as the depth of closure shifts seawards with increasing foreshore nourishment. Besides the dominant variables, a correlation was found between the Bma_y, horizontal movement of the dune foot, increase in dune volume and active dune width. A larger increase in dune volume and horizontal movement of the dune foot, correlates with a lower Bma_y relative to NAP. With almost no increase in dune volume and horizontal movement of the dune foot, the Bma_y lies higher relative to NAP.

- **What are the recommendations to incorporate the understanding of long term changes in future strategical plans and policies ?**

  Nourishment has been an increasing approach in maintaining the coast. This research shows that nourishment not only positively affects the characteristics of the coast, but also changes its future dynamic equilibrium. The existing Bruun rule is used as the basic fundamental adaption of the coastline profile, even though its confined approach. The results of this research show, where the future adaption upon the existing Bruun rule should be improved, for the use in the complex coastal zone. Clusters originate from known and unknown processes that influence their unique characteristics. Each specific cluster requires a unique strategical plan and policy for the future preservation of the coastline profile of the management area of HHNK.

V.2.2 *Main research question*

Combining the results of the previous research questions, the following main research question is answered.

> *What is the expected future coastline profile change of the Hoogheemraadschap Hollands Noorderkwartier area based on historical measurements ?*

There are hardly any areas that show a retreat in dune volume or foreshore volume. On average most clusters have an increasing volume and positive changes due to nourishment. Resulting in human-induced nourishment, dominating changes along the coastline for most clusters. This is more strongly seen in the foreshore rather than the dunes. In the dunes, the boundary between marine and aeolian transport seems to play a role in the positive changes. As there are exceptions for these general statements, each cluster has their own unique changes, dominated by their specific processes. While the exact processes remain unknown, the locations and changes are identified by the clusters. Following the trend of increasing nourishment in the recent years, it's safe to assume this policy will carry on. Nourishment results in positive changes, but also in changes of the dynamic equilibrium of the coastline profile.

V.3 Recommendations

V.3.1 Application

The main focus of this research, is to give more insight into the complex coastal zone of the management area of HHNK. The clusters are an ideal input, to be interpreted by HHNK, for future strategical plans and policies from a new perspective. Making it also possible to add future new information on top of the clusters, to find new patterns and insights. The current nourishment approach positively affects the coastline. The clusters that had lesser nourishment contribution, should be included in the future strategical plans and policies. The clusters that show anomalies, like loss of volume in the dunes or loss of volume in the foreshore, should be researched for underlying causes and for future prevention of recession.

V.3.2 Further research

This research has given more insight in the variables and their underlying correlations. Due to the time restriction of this research, more research is required to substantiate the effects on the long term changes. As referred to in the discussion, two of the main interesting results come from the depth of closure and the boundary between marine and aeolian transport.

Depth of Closure and Nourishment
Better understanding of the depth of closure would be beneficial for HHNK, as the amount of nourishment is calculated based on the boundaries, such as the inner and outer depth of closure Rijkswaterstaat [2020]. This nourishment affects the long-term changes of the coastline profile. With main-cluster 2A representing a strong correlation between the shifting depth of closure and the amount of nourishment. These results compare to lab experiments Aktinson and Baldock [2020]. If time wasn't a restriction in this research, the same experimental set-up could be used to answer the remaining questions;

As seen in the Bruun rule, the depth of closure remains at the same location in time. If the depth of closure changes, how would this affect the future equilibrium of the profile ? Is this relationship linear ? Does this increase the active profile width and therefore require more sediment ? These are a handful of questions which could be the basis for an experimental setup. Having more insight in the depth of closure, could also help understand the ecological effects of nourishment, as this has been becoming a more popular research subject Rijkswaterstaat [2020]. Where the area below the depth of closure changes from a static zone into a dynamic zone, it affects the habitat of animal species.

Boundary between marine and aeolian transport y-coordinate
The boundary between marine and aeolian transport, shows an interesting correlation between the derived variables located in the dune area. This correlations seems to be less cluster dependent. The causation between these variables couldn't be substantiated in this research. Deeper understanding of both the derivation, with the use of the standard deviation method and physical understanding, could provide more insight in the changes in the aeolian transport area.

A. Appendix - Sketch/location variables

Appendix A.1 represents the derived variables which have been motioned in chapter III.4.1. Every derived variable is represented by a symbol, as seen in figure A.1. The legend shows widths, gradients and volumes by combining two symbols, representing a clear overview of the various variables which are derived from the cross-shore profile of the transects.

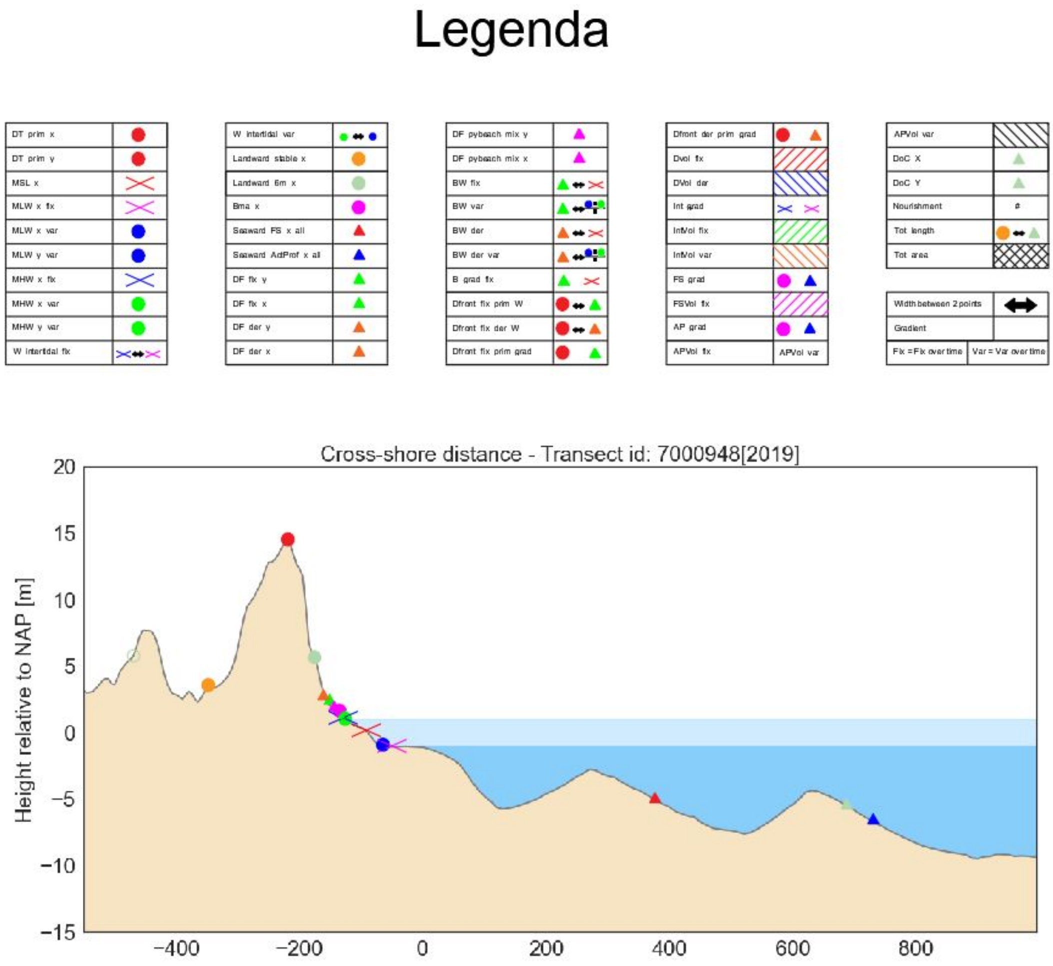


Figure A.1: Locations of every derived variable for transect 7000984 [2019].

B. Appendix - Background variables

Appendix B introduces a more detailed explanation and background of each individual variable. As the data set contains a great number of variables, this chapter focuses on the variable used in the primary data set of this research.

There are various methods of deriving the dune foot as seen in figure A.1. This research will use the dune foot from a prior data set, denoted as dune foot from the new data set [7].

B.0.1 *Dune Crest*

The dune crest is defined as the highest location of the dunes, as seen in figure B.1. The y-coordinate of this variable is denoted as *DT_prim_y* and it's x-coordinate is denoted as *DT_prim_x*. The derivation of this variable uses the scipy.signal.find_peaks module [8]. This module is based on the prominence of a peak. This is a measure to calculate how much a peak stands out relative to its surrounding. The parameters used are a height of 5 and prominence of 2.0.

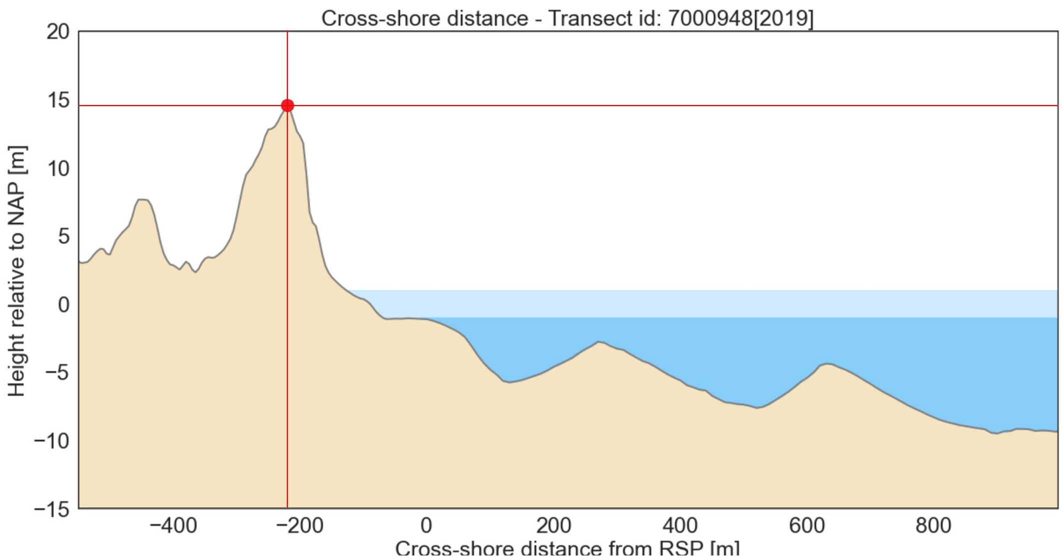


Figure B.1: Derived variable: Dune Top *[DT_prim_y] and [DT_prim_x]*.

B.0.2 *Intertidal width*

The intertidal width is defined as the horizontal distance between the Mean Low Water level and the Mean High Water level, as seen in figure B.2. The intertidal width is denoted as *W_intertidal_var*. The Mean High Water level is taken as a variable location which is included in the JarKus database. This variable height differs in space, from north to south, but is steady in time. The variable values lie close to the fixed locations. The fixed y-coordinate for the Mean High Water level is located at +1 m NAP. For the Mean Low Water level, the y-coordinate is located at -1 m NAP.

[7] http://opendap.deltares.nl/thredds/catalog/opendap/rijkswaterstaat/DuneFoot/catalog.html?dataset=varopendap/rijkswaterstaat/DuneFoot/DF_2nd_deriv.nc

[8] https://docs.scipy.org/doc/scipy/reference/generated/scipy.signal.find_peaks.html

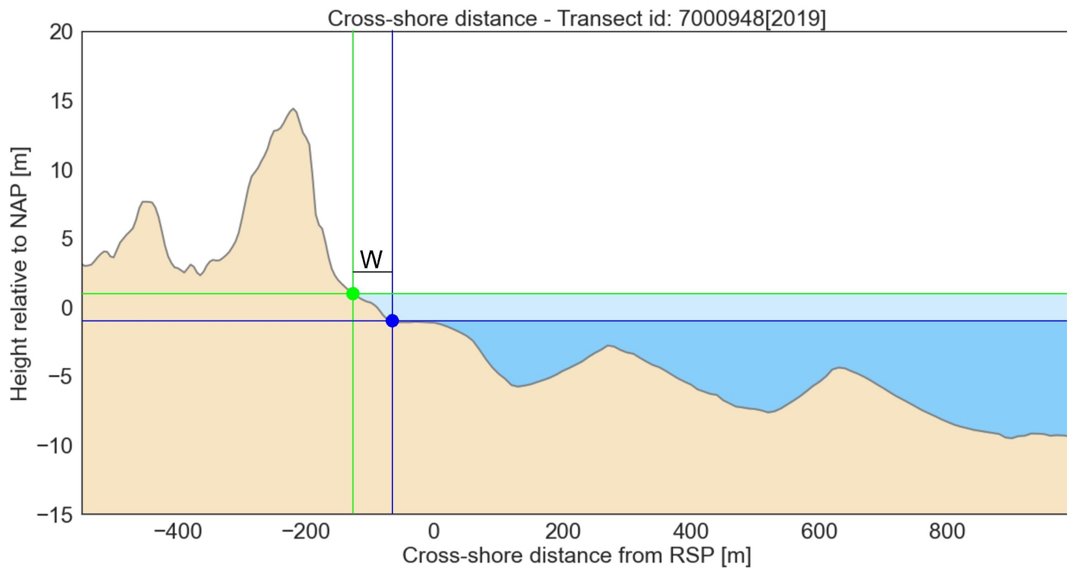


Figure B.2: Derived variable: Intertidal width *[W_intertidal_var]*.

B.0.3 Beach Width

The beach width is defined as the horizontal distance between the Mean Sea Level and the Dune Foot, as seen in figure B.3. The y-coordinate of the Mean Sea Level is the mean of the Mean High Water level and Mean Low Water level. For the fixed location this is located at 0m NAP. The beach width can include both a fixed Mean Sea Level and Dune Foot. The chosen width includes the fixed Mean Sea Level and both the Dune Foot from the new data set, denoted as *beachwidthvar2*. The beach width including a fixed Dune Foot is denoted as *BW_var*, which includes a fixed Dune Foot which is located at +3m NAP.

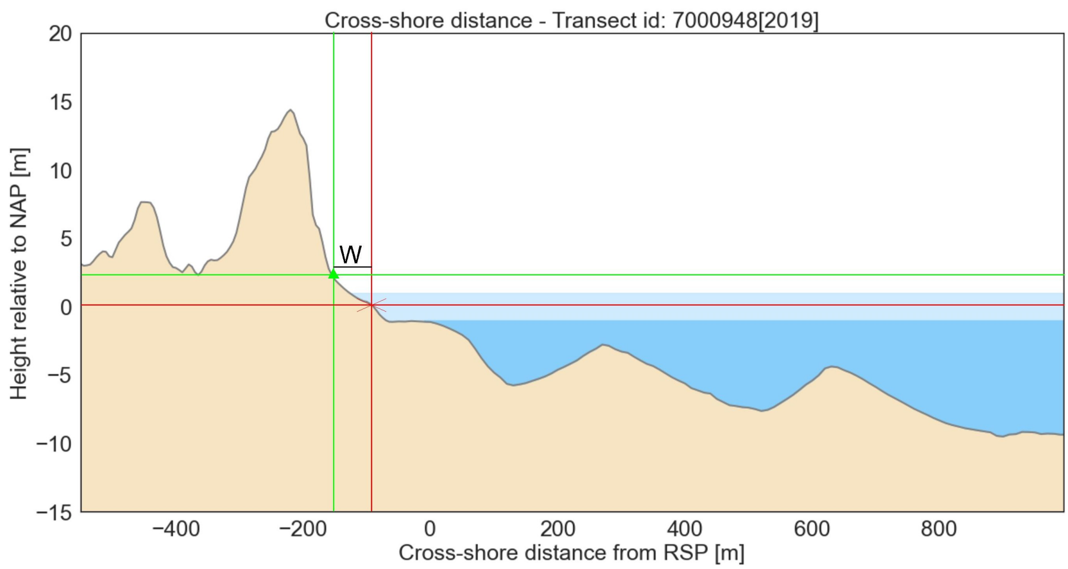


Figure B.3: Derived variable: Beach Width *[beachwidthvar2]*.

B.0.4 *Beach Gradient*

The beach gradient is defined as the gradient between the Mean Sea Level and the Dune Foot, as seen in figure B.4. The beach gradient used in this research is denoted as *B_grad_fix*. This gradient is defined by the fixed beach width.

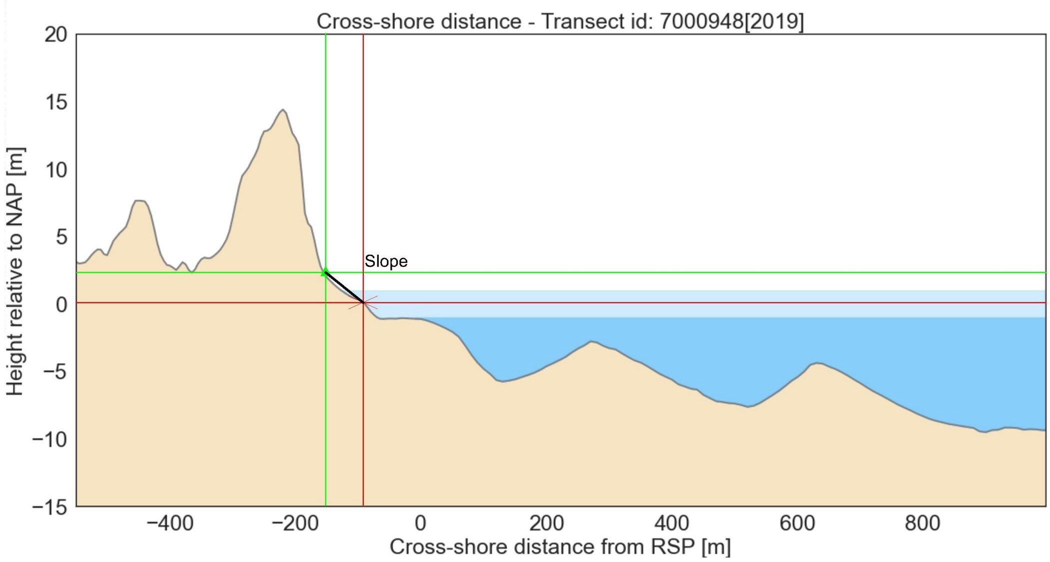


Figure B.4: Derived variable: Beach Gradient *[B_grad_fix]*.

B.0.5 *Boundary between Marine and Aeolian transport*

The boundary between marine and aeolian transport is located by isolating the variance along the coast De Vries et al. [2010]. This variable is denoted as *Bma_x*.

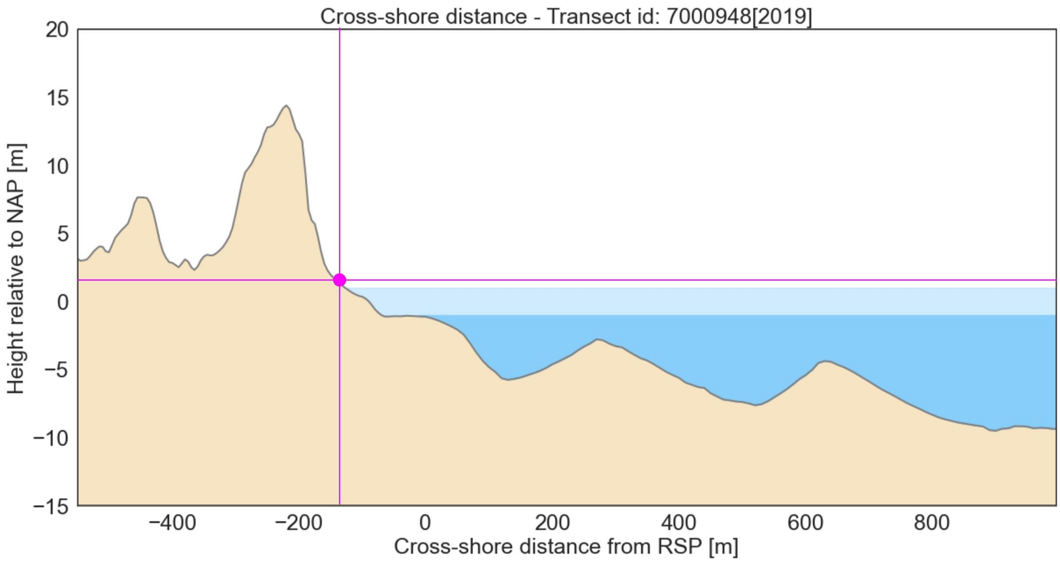


Figure B.5: Derived variable: Boundary between marine and aeolian transport *[Bma]*.

B.0.6 Foreshore Gradient

The foreshore gradient is defined as the gradient between the marine and aeolian transport boundary and the fixed depth of closure, as seen in figure B.6. The fixed depth of closure, denoted as *Seaward_ActProf_x*, is located at -8m NAP. The foreshore gradient is denoted as *FS_grad*.

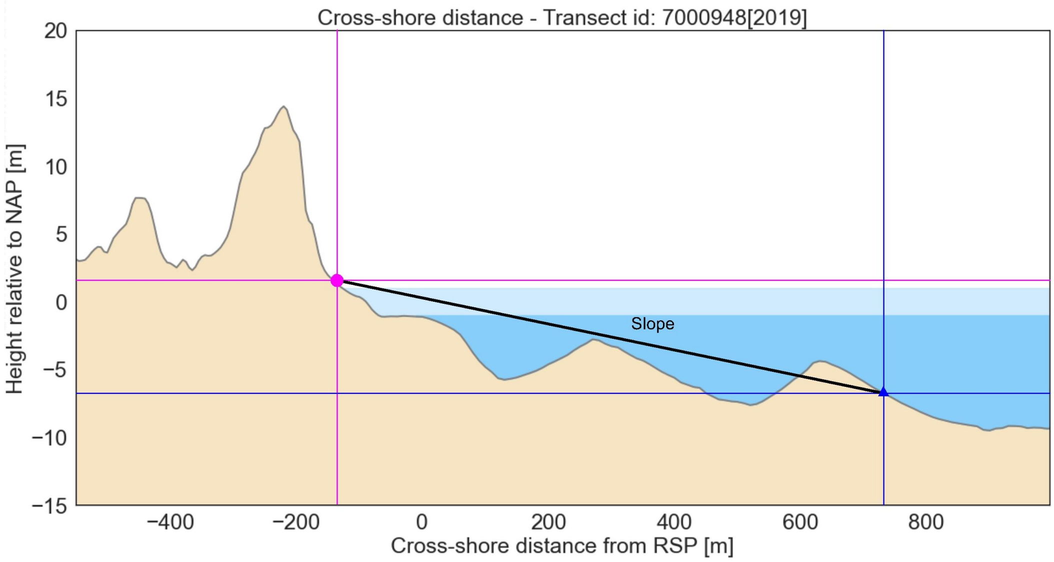


Figure B.6: Derived variable: Foreshore Gradient *[FS_grad]*.

B.0.7 Dune Foot

There have been various methods of deriving the dune foot, as seen in figure B.7. This research uses the dune foot from the new data set, with the y-coordinate denoted by *DF_y2* and the x-coordinate denoted by *DF_x2*.

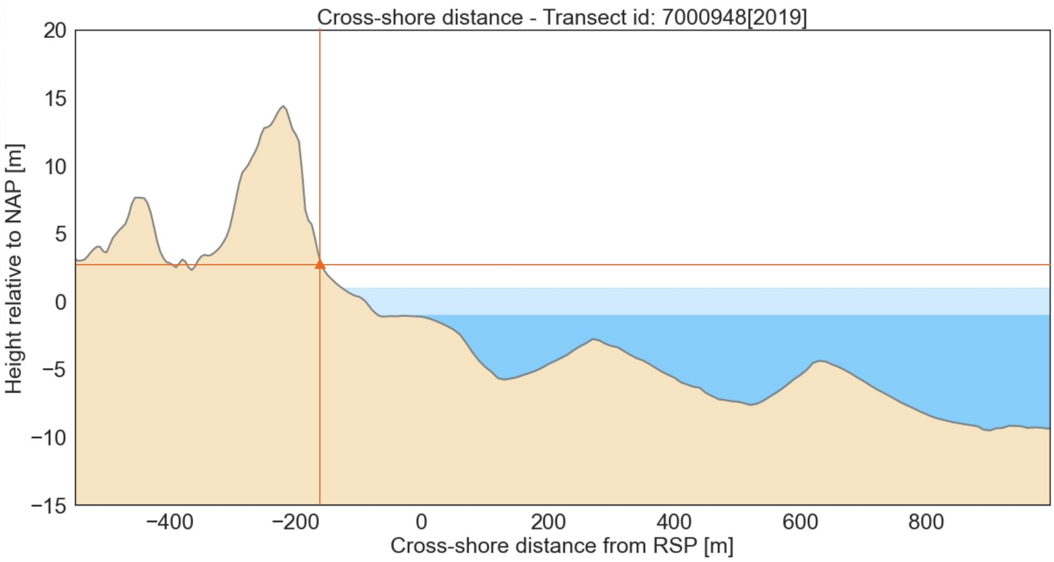


Figure B.7: Derived variable: Dune Foot *[DF_x]* and *[DF_y]*.

B.0.8 *Dune Front*

The dune front is defined as the width between the dune crest and the dune foot from the new data set, as seen in figure B.8. The variable is denoted as *dunefront2*.

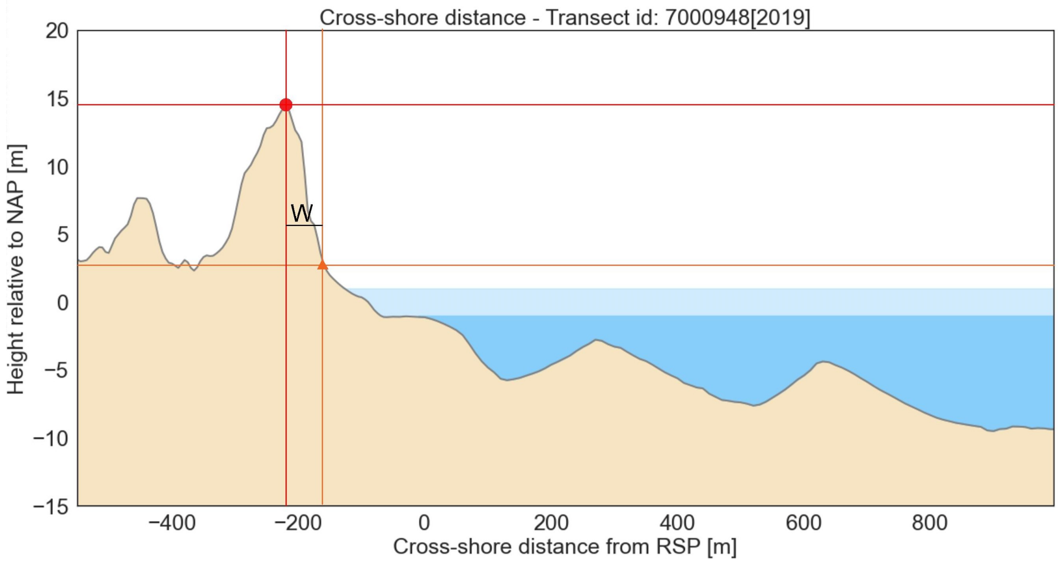


Figure B.8: Derived variable: Dune front *[dunefront2]*.

B.0.9 *Dune Gradient*

The dune gradient is defined as the gradient of the dune front, as seen in figure B.9. This variable is denoted as *DFront_der_prim_grad*. The dune front is defined by the dune crest and the dune foot of the new data set.

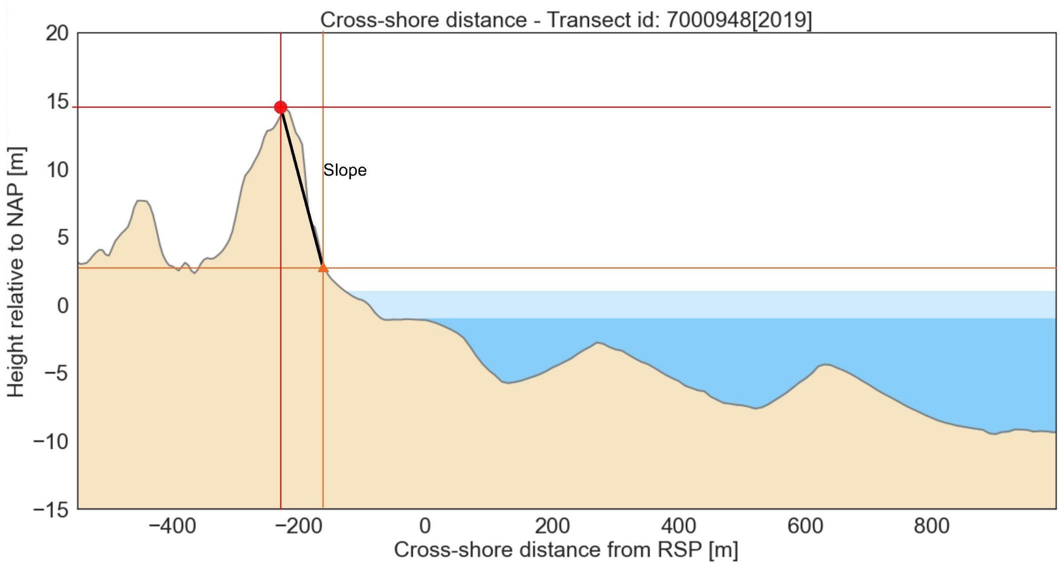


Figure B.9: Derived variable: Dune Gradient *[DFront_der_prim_grad]*.

B.0.10 Landward 6 meter

The landward 6 meter location, is the first location at +6 meter NAP going from the dune crest seawards, as seen in figure B.10. This variable is denoted as *Landward_6m*.

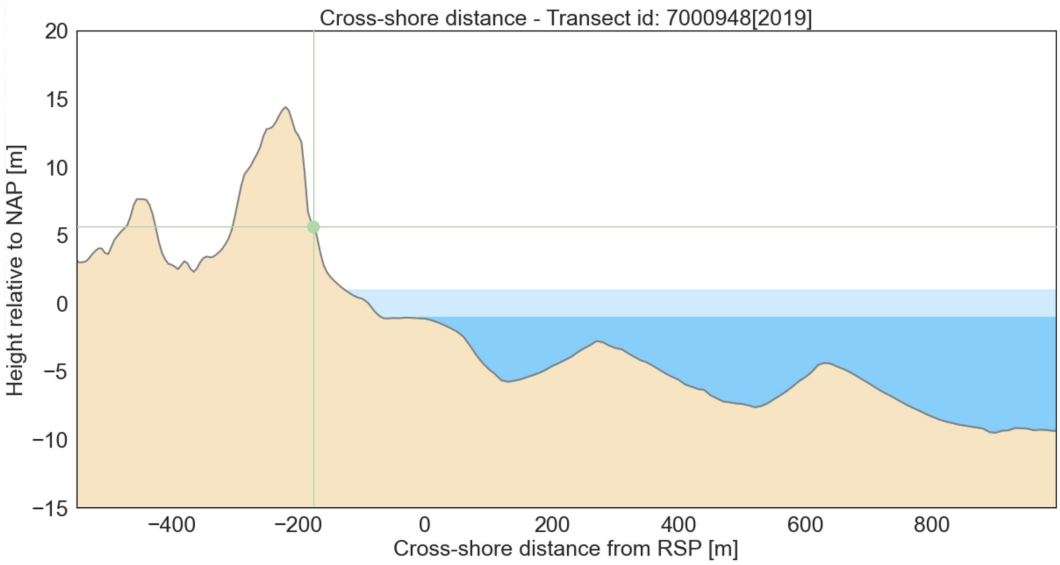


Figure B.10: Derived variable: Landward 6 meter *[Landward_6m]*.

B.0.11 Seaward Foreshore

The seaward foreshore location is located at the y-coordinate of -4m NAP, as seen in figure B.11. The variable is denoted as *seaward_FS_x_all*.

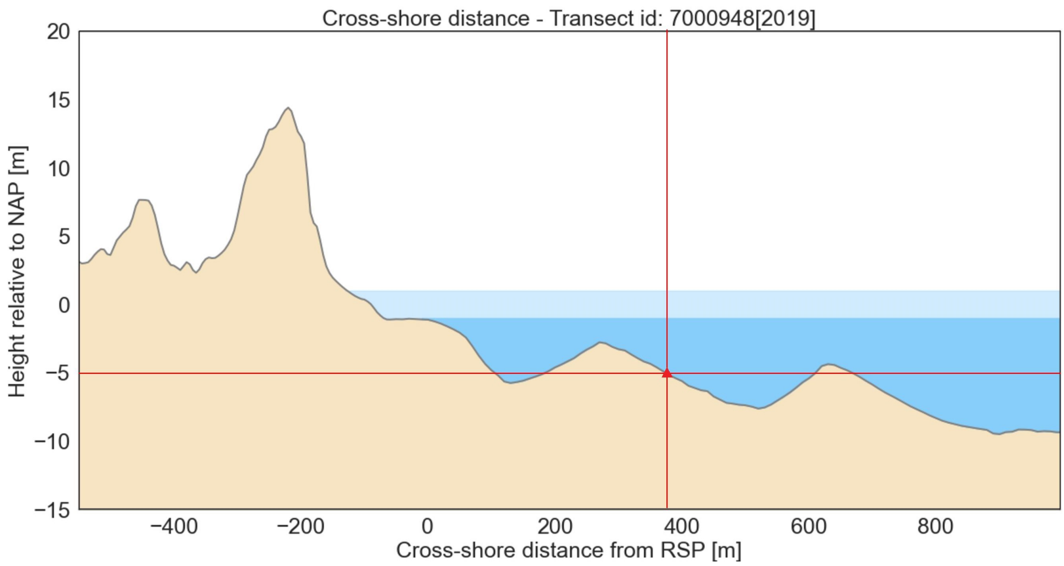


Figure B.11: Derived variable: Seaward Foreshore *[seaward_FS_x_all]*.

B.0.12 *Landward Point*

The landward point is defined as the first location which crosses the standard deviation threshold, going landward from the dune crest, as seen in figure B.12. This location is denoted as *beachy*

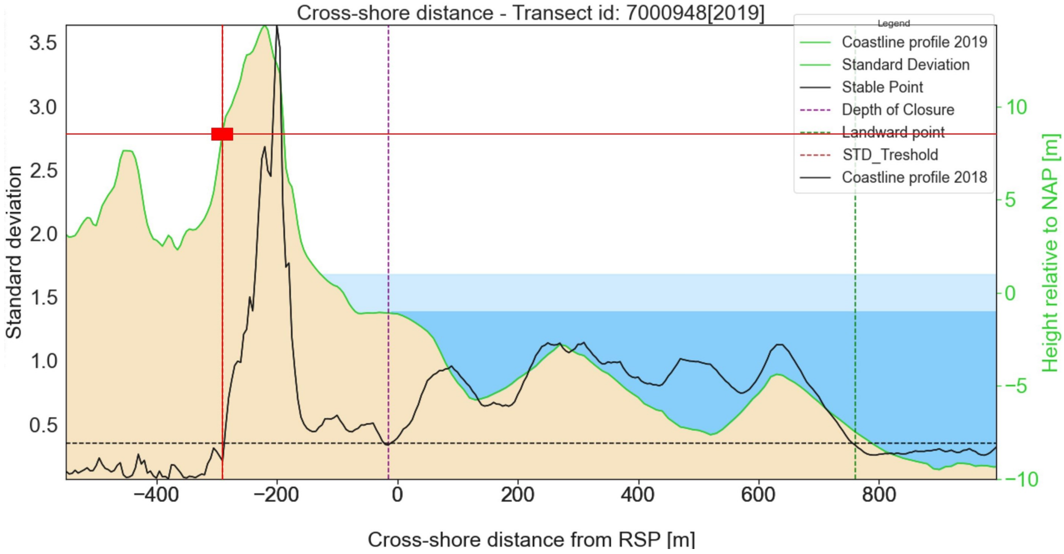


Figure B.12: Derived variable: Landward point *[beachy]*.

B.0.13 *Bma_y*

The y-coordinate of the Boundary between marine and aeolian transport is defined at the minimum location between the peaks of the standard deviation method. The first peak is located at the dunes and the second peak is located in the foreshore, as seen in figure B.13. This location is denoted as *Bma_y*, which represents the location with the least amount of variability.

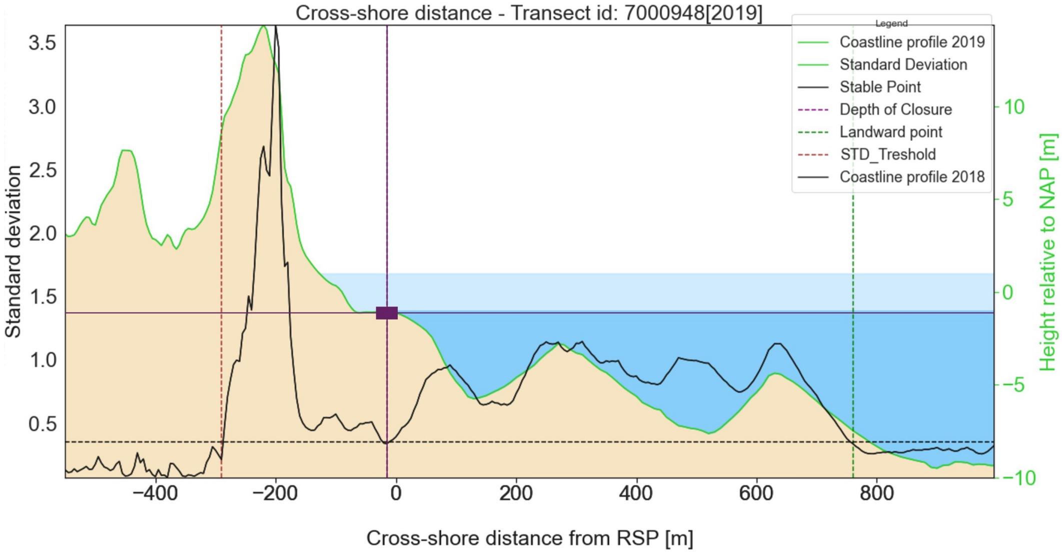


Figure B.13: Derived variable: Stable point *[Bma_y]*.

B.0.14 *Depth of Closure*

The depth of closure is defined as the location which shows almost no variability over the years. This location is derived by defining the crossing of the standard deviation with the threshold of 0.25 - 0.35 as defined by II.6. This variable is denoted as *docy*.

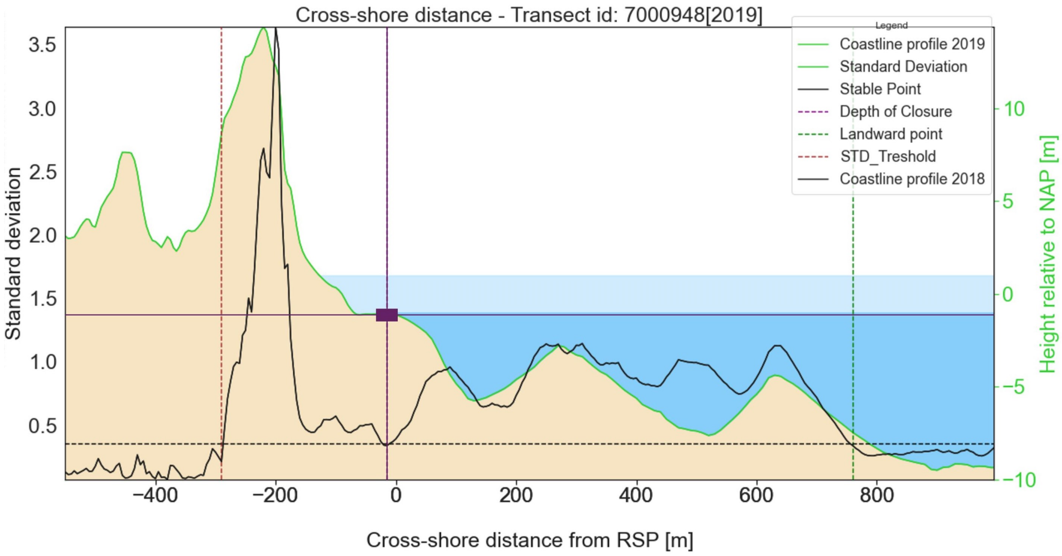


Figure B.14: Derived variable: Depth of Closure *[docy]*.

B.0.15 *Active Width Dune*

The active width of the dunes, is the width between the landward point and the stable point, as seen in figure B.15. This width is denoted as B.15.

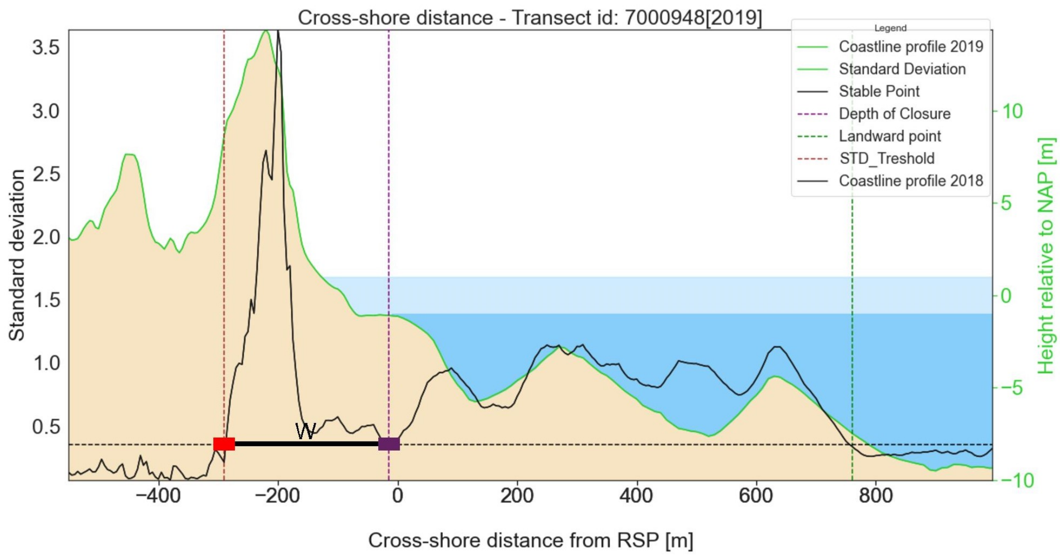


Figure B.15: Derived variable: Active width Dune *[Dune_W]*.

B.0.16 *Active Width Foreshore*

The active width of the foreshore, is the width between the stable point and the depth of closure, as seen in figure B.16. This variable is denoted as *FS_W*.

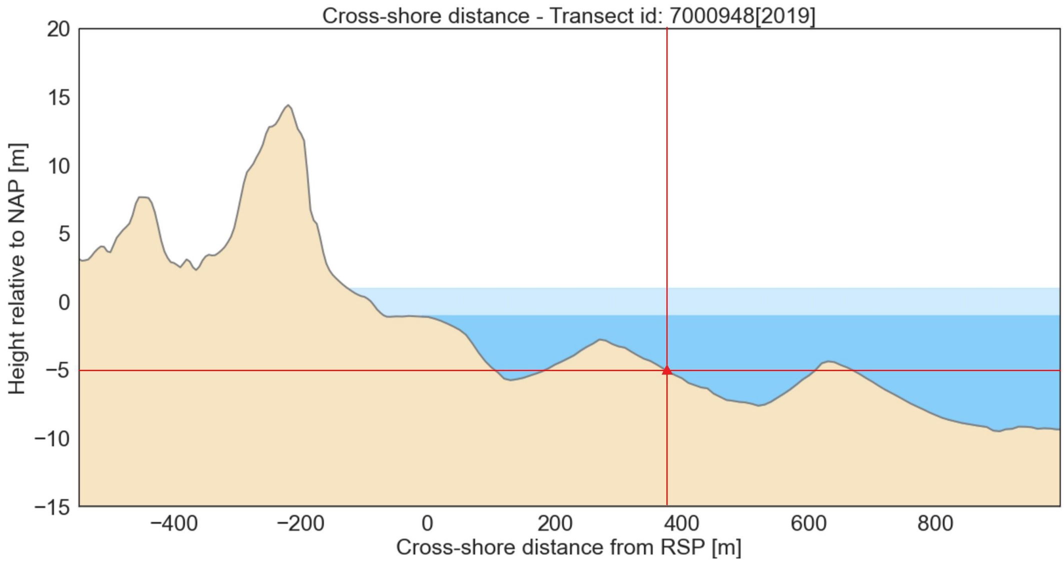


Figure B.16: Derived variable: Active width Foreshore *[FS_W]*.

B.0.17 *Total Active Profile Width*

The total active profile width is defined by the sum of the active profile of the dunes and the active profile of the foreshore. Meaning the width between the landward point and the depth of closure as seen in figure B.17. This indicates that in the years measured, almost all variability of the profile can be seen between these two points.

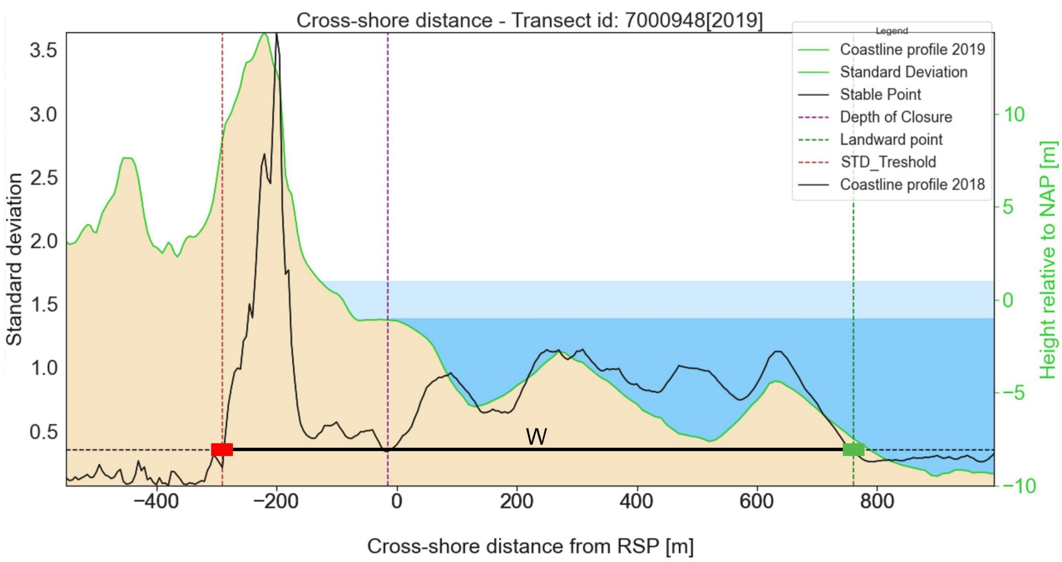


Figure B.17: Derived variable: Total active profile width *[ActiveProfile_W]*.

B.0.18 *Dune Volume*

The dune volume is defined as the volume between the landward point and the stable point, as seen in the shaded area in figure B.18. The actual volume will depend on the chosen y-coordinate, therefore this research looks at the change in volume in this area. With its variable denoted by *dunesm*.

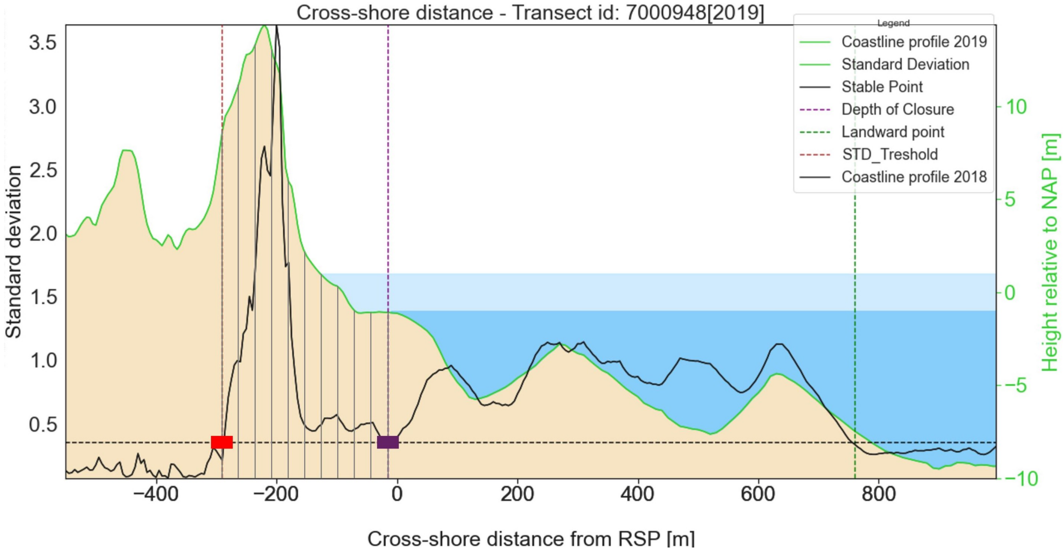


Figure B.18: Derived variable: Dune volume *[dunesm]*.

B.0.19 *Foreshore Volume*

The foreshore volume is defined as the volume between the stable point and the depth of closure, as seen in the shaded area in figure B.19. The actual volume will depend on the chosen y-coordinate, therefore this research looks at the change in volume in this area. With its variable denoted by *foreshorem*.

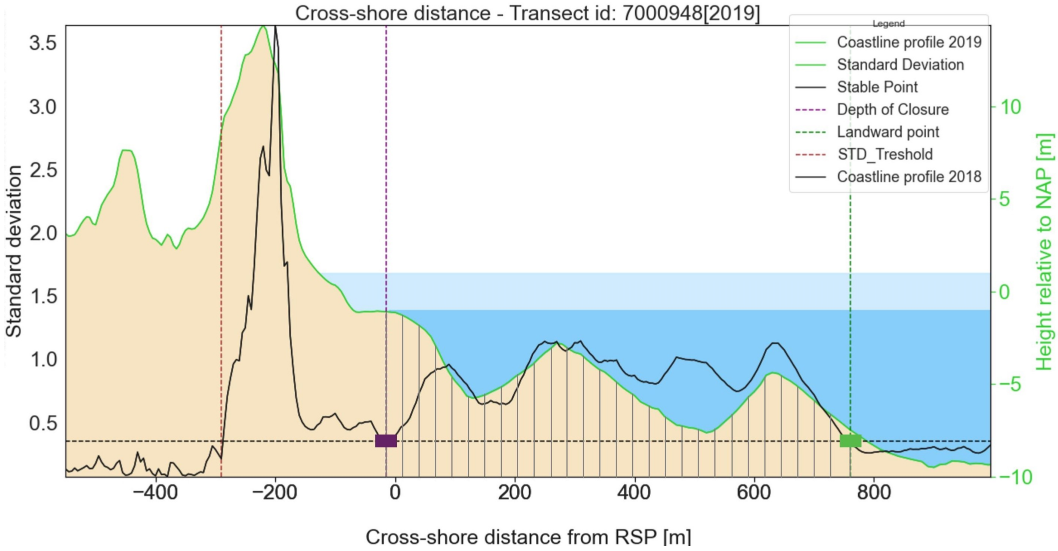


Figure B.19: Derived variable: Foreshore volume *[foreshorem]*.

C. Appendix - Component Planes Variables

Appendix C represents the component planes of each individual variable used in the different data sets. The component planes are a fast method to find correlations between different variables and indicate the range of values. The component planes of the first initial data set is represented in Figure C.1. The component planes of the second initial data set is represented by figure C.2. Finally, the component planes represented by the finalized data set can be seen in figure C.3.

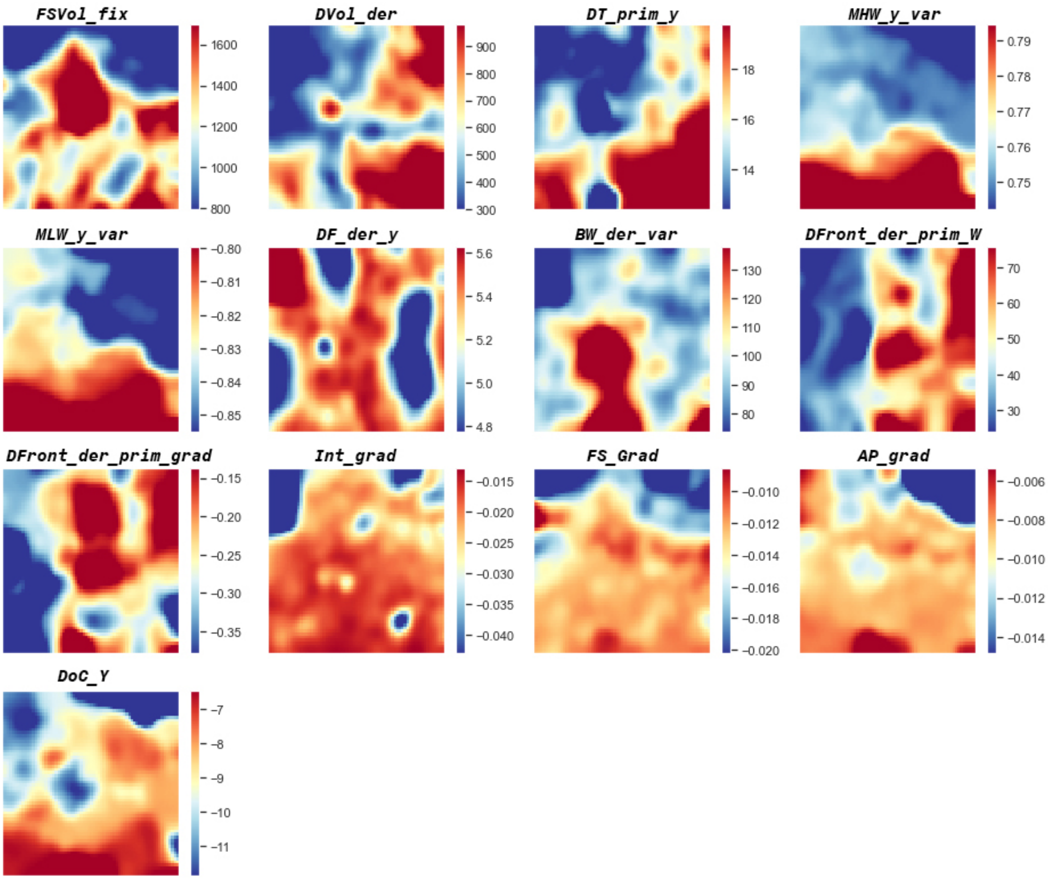


Figure C.1: All component planes for the initial data set 1A.

.

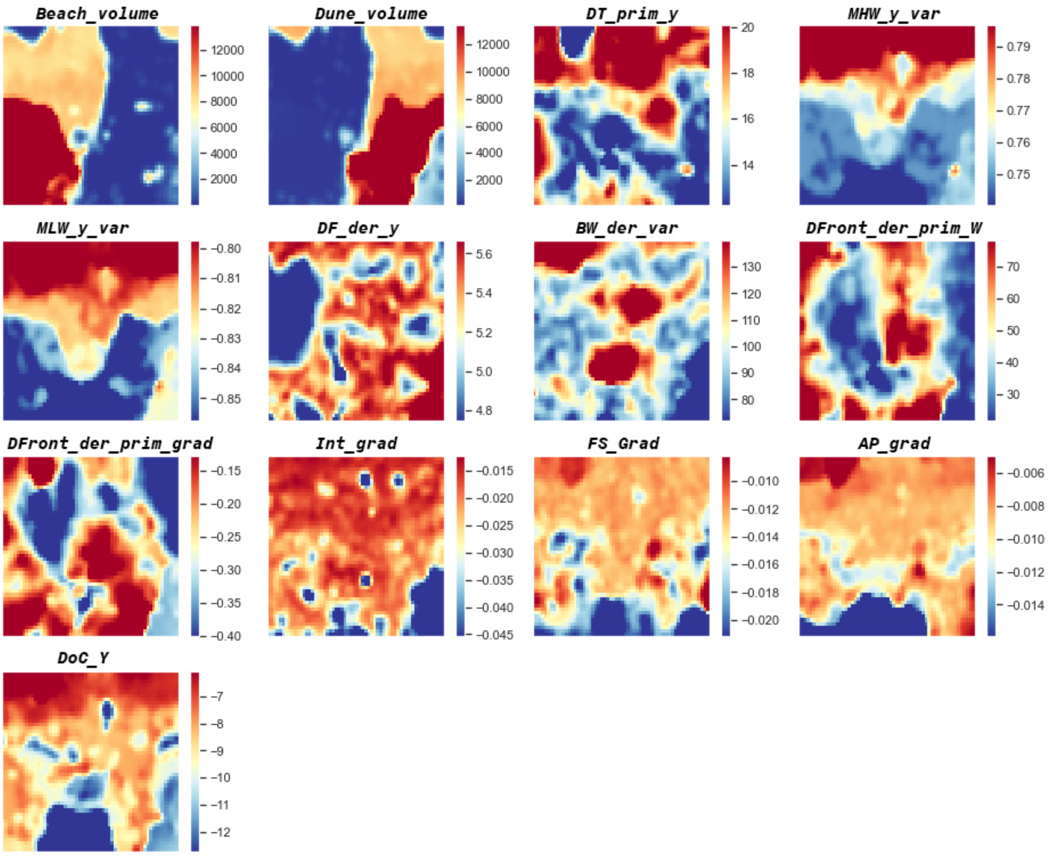


Figure C.2: All component planes for the initial data set 1B.

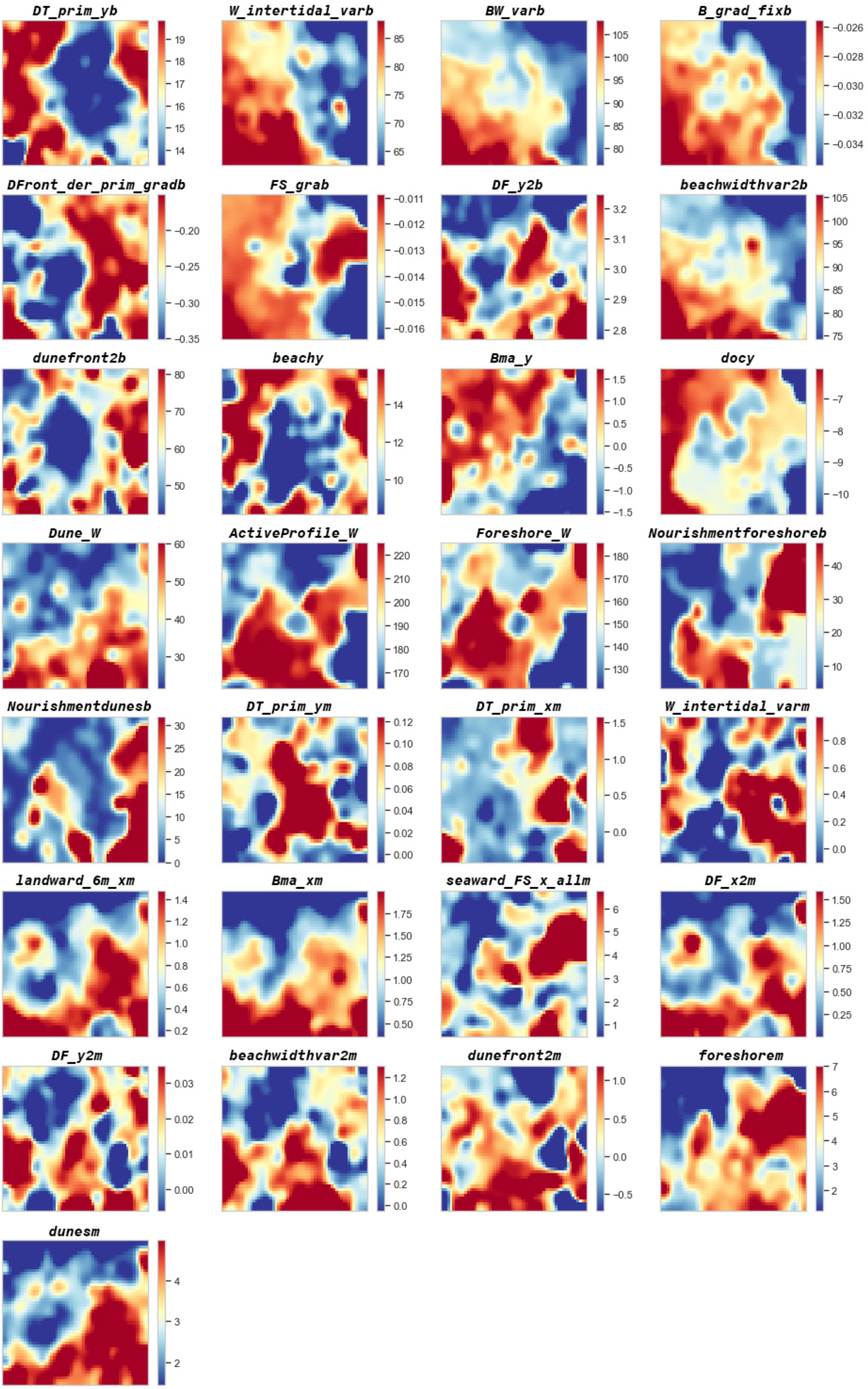


Figure C.3: All component planes for the final data set.

D. Appendix - Clustered Components Planes

Appendix D represents the clustered component planes of each individual variable. The clustered component planes represent the clusters in the output space, by representing them back onto the input space. This is done by plotting the contours of the clusters back on each individual component plane. The clustered component planes of the first initial data set is represented in Figure D.1. The clustered component planes of the second initial data set is represented by figure D.2. Finally the clustered component planes represented by the finalized data set can be seen in figure D.3.

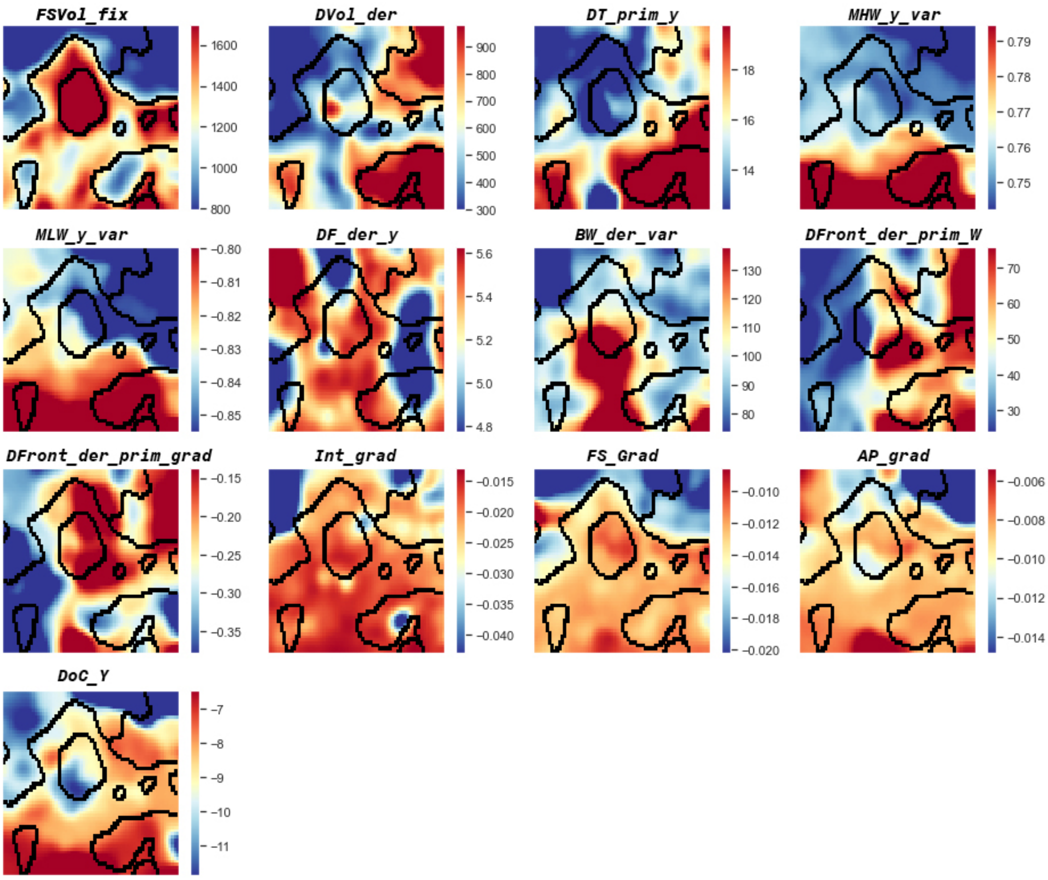


Figure D.1: All clustered component planes for the initial data set 1A.

.

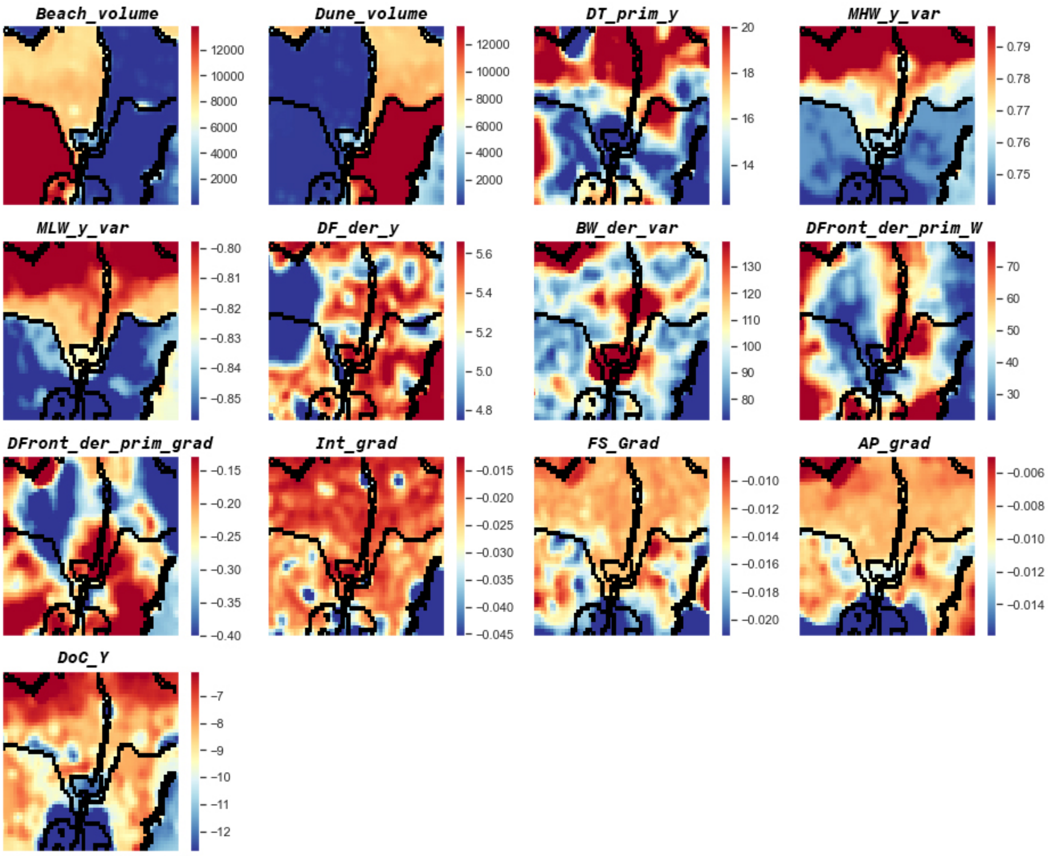


Figure D.2: All clustered component planes for the initial data set 1B.

Figure D.3: All clustered component planes for the final data set.

E. Appendix - Quantitative Characteristics Clusters

Appendix E represents the quantitative characteristics of both the main-clusters and sub-clusters. Each variable contains it's 25th percentile and 75th percentile. The bold results represent the average value of each cluster. The characteristic slope of each individual variable is represented in table 10. The characteristic starting value of each cluster is represented in table 11.

Table 10: All characteristic changes of the variables for both the main-clusters and the sub-clusters.

Mean/Clusters	0	1	2A	2B	3	4A	4B	5A	5B
25%	0.004	-0.004	0.024	0.011	-0.040	0.013	0.062	0.003	0.011
DT_prim_ym [NAP/year]	**0.013**	**0.005**	**0.069**	**0.046**	**-0.036**	**0.060**	**0.127**	**0.038**	**0.025**
75%	0.005	0.029	0.121	0.048	0.007	0.089	0.171	0.085	0.044
25%	-0.012	0.00	-0.15	0.01	0.03	-0.03	-0.07	-0.01	0.01
DT_prim_xm [m/year]	**-0.32**	**0.54**	**0.35**	**0.12**	**2.27**	**0.37**	**0.41**	**1.46**	**1.62**
75%	-0.04	0.45	0.50	0.32	4.53	0.86	0.63	1.79	2.44
25%	0.16	-0.17	-0.40	-0.08	-0.09	0.05	0.97	0.89	0.92
W_intertidal_varm [m/year]	**0.38**	**0.23**	**0.17**	**0.60**	**0.23**	**0.45**	**1.28**	**1.03**	**0.98**
75%	0.58	0.58	0.62	1.24	0.60	0.91	1.68	1.22	1.03
25%	0.29	0.43	0.58	0.42	1.57	-0.10	1.28	1.00	-0.36
landward_6m_xm [m/year]	**0.29**	**0.77**	**0.95**	**1.01**	**2.45**	**0.23**	**1.43**	**1.50**	**-0.35**
75%	0.29	1.09	1.43	1.52	3.07	0.53	1.62	1.81	-0.33
25%	1.50	0.88	1.02	0.06	2.57	-0.03	1.46	1.17	-0.35
Bma_xm [m/year]	**2.00**	**1.10**	**1.56**	**1.28**	**3.14**	**0.38**	**1.51**	**1.73**	**-0.32**
75%	2.51	1.30	2.16	2.21	3.28	0.86	1.60	2.46	-0.31
25%	8.72	8.60	1.25	0.09	2.61	0.34	0.46	2.02	2.15
seaward_FS_x_allm [m/year]	**9.32**	**9.70**	**3.70**	**2.09**	**3.95**	**1.95**	**2.30**	**2.41**	**2.66**
75%	11.56	12.61	5.09	3.67	5.72	3.42	3.99	3.63	3.06
25%	0.15	0.23	0.52	0.36	1.12	-0.23	1.41	0.98	-0.66
DF_x2m [m/year]	**0.34**	**0.55**	**1.06**	**1.55**	**2.58**	**0.12**	**1.70**	**1.51**	**-0.64**
75%	0.49	0.90	1.52	2.30	3.26	0.48	1.85	2.20	-0.63
25%	0.01	0.01	0.00	0.00	0.00	0.00	-0.02	0.00	0.02
DF_y2m [NAP/year]	**0.03**	**0.03**	**0.02**	**0.00**	**0.01**	**0.01**	**0.00**	**0.02**	**0.02**
75%	0.04	0.04	0.03	0.01	0.03	0.02	0.01	0.03	0.01
25%	2.32	0.57	0.33	-0.31	0.78	-0.04	-0.04	0.24	0.75
beachwidthvar2m [m/year]	**2.67**	**0.83**	**0.69**	**-0.10**	**1.31**	**0.45**	**0.36**	**0.42**	**0.76**
75%	3.25	1.15	1.19	0.31	1.77	0.77	1.12	0.66	0.77
25%	0.36	0.15	0.12	0.49	-0.53	-0.41	0.54	-0.82	-2.85
dunefront2m [m/year]	**0.87**	**0.21**	**0.52**	**0.61**	**0.49**	**-0.13**	**1.01**	**0.08**	**-2.04**
75%	0.84	0.92	1.08	0.65	1.39	0.36	1.80	1.55	-0.44
25%	6.96	7.58	4.27	-12.21	2.23	0.23	2.38	4.28	1.58
foreshorem $[m^3/year]$	**6.64**	**8.55**	**5.33**	**-6.20**	**4.05**	**1.63**	**3.85**	**5.01**	**1.98**
75%	7.26	10.52	6.41	-3.13	5.37	2.97	5.01	5.80	2.58
25%	3.83	1.61	2.23	1.49	3.53	0.62	4.48	4.37	0.20
dunesm $[m^3/year]$	**5.13**	**2.99**	**3.69**	**3.56**	**4.53**	**1.51**	**5.18**	**5.56**	**0.35**
75%	6.14	3.82	5.19	5.91	6.20	2.43	5.95	6.92	0.59
25%	2.6	-0.1	-0.5	1.0	-1.6	-5.0	-2.7	-2.1	-12
DFront_der_ [m] prim_gradm $[\times 10^{-3}]$	**6.7**	**1.0**	**3.4**	**2.8**	**0.00**	**-1.4**	**0.3**	**0.8**	**-8.7**
75%	10	3.5	8.1	2.6	1.9	2.0	4.1	3.2	-2.9
25%	0.46	0.19	0.08	-0.55	0.00	0.02	-0.22	-0.33	-0.24
B_grad_fixm $[\times 10^{-3}]$	**0.52**	**0.33**	**0.25**	**-0.36**	**0.02**	**0.13**	**-0.04**	**-0.21**	**0.32**
75%	0.55	0.43	0.42	-0.22	0.06	0.22	0.11	0.07	0.44

Table 11: All characteristic values of the variables for both the main-clusters and the sub-clusters.

Mean/Clusters	0	1	2A	2B	3	4A	4B	5A	5B
# of profiles	[5]	[22]	[51]	[5]	[12]	[77]	[8]	[19]	[3]
25%	14.0	18.5	13.6	15.5	9.4	13.1	13.8	14.8	14.6
DT_prim_yb [NAP]	**17.3**	**19.2**	**16.3**	**16.9**	**15.1**	**17.5**	**14.9**	**16.1**	**15.8**
75%	19.0	20.5	18.8	17.7	20.9	20.8	15.8	17.6	16.4
25%	74	63	76	44	102	73	64	49	55
W_intertidal_varb [m]	**77**	**65**	**82**	**46**	**107**	**77**	**67**	**53**	**57**
75%	82	68	90	48	109	82	69	57	59
25%	102	77	90	68	145	81	84	74	65
beachwidthvar2b [m]	**102**	**77**	**97**	**70**	**149**	**85**	**88**	**78**	**66**
75%	107	81	101	74	157	91	93	81	66
25%	-0.028	-0.039	-0.029	-0.038	-0.017	-0.032	-0.032	-0.039	-0.039
B_grad_fixb [-]	**-0.029**	**-0.037**	**-0.028**	**-0.037**	**-0.017**	**-0.030**	**-0.030**	**-0.037**	**-0.038**
75%	-0.027	-0.035	-0.027	-0.036	-0.016	-0.028	-0.028	-0.034	-0.037
25%	-0.431	-0.235	-0.338	-0.257	-0.133	-0.408	-0.226	-0.210	-0.238
DFront_der_ prim_gradb [-]	**-0.363**	**-0.208**	**-0.275**	**-0.209**	**-0.106**	**-0.315**	**-0.191**	**-0.152**	**-0.216**
75%	-0.348	-0.171	-0.188	-0.143	-0.057	-0.223	-0.157	-0.084	-0.179
25%	-0.015	-0.012	-0.013	-0.028	-0.011	-0.014	-0.016	-0.021	-0.018
FS_gradb [-]	**-0.015**	**-0.012**	**-0.012**	**-0.024**	**-0.010**	**-0.013**	**-0.016**	**-0.020**	**-0.017**
75%	-0.015	-0.012	-0.013	-0.028	-0.011	-0.014	-0.016	-0.021	-0.018
25%	2.9	2.8	2.8	2.9	3.1	2.8	3.1	3.0	2.7
DF_y2b [NAP]	**2.9**	**2.9**	**3.0**	**3.0**	**3.3**	**2.9**	**3.1**	**3.2**	**2.8**
75%	3.0	3.1	3.1	3.1	3.5	3.0	3.2	3.5	2.8
25%	28	68	43	51	70	41	52	56	53
dunefront2b [m]	**50**	**76**	**55**	**65**	**89**	**53**	**63**	**94**	**75**
75%	70	83	60	71	84	64	73	125	87
25%	9.2	13.3	7.6	12.5	8.0	9.8	8.8	11.5	11.9
beachy [NAP]	**10.6**	**15.6**	**10.5**	**14.2**	**12.0**	**13.1**	**10.2**	**13.6**	**13.5**
75%	13.5	18.7	12.7	15.8	15.4	16.6	11.9	15.4	14.8
25%	-5.8	-1.2	-1.0	-4.0	-0.6	0.8	-0.8	-2.3	0
Bma_y [NAP]	**-5.6**	**-0.6**	**0**	**-2.8**	**1.4**	**1.1**	**0.1**	**-1.6**	**0.4**
75%	-5.4	-0.4	1.1	-1.7	2.0	1.6	1.3	-1.6	1.0
25%	-10.4	-8.9	-8.6	-19.8	-5.4	-6.9	-8.3	-11.9	-6.9
docy [NAP]	**-10.1**	**-8.9**	**-8.6**	**-19.8**	**-5.4**	**-6.9**	**-8.3**	**-11.9**	**-6.9**
75%	-9.8	-7.9	-8.2	-19.7	-4.9	-6.0	-7.6	-9.1	-6.6
25%	710	163	158	260	155	110	155	225	115
Dune_W [m]	**713**	**192**	**213**	**254**	**232**	**144**	**222**	**257**	**135**
75%	785	214	263	275	326	160	284	295	150
25%	1100	986	1103	1445	881	815	794	630	618
ActiveProfile_W [m]	**1100**	**1016**	**1130**	**1554**	**1000**	**863**	**869**	**689**	**643**
75%	1155	1049	1140	1595	1096	915	959	720	673
25%	370	820	873	1170	593	670	556	348	490
Foreshore_W [m]	**387**	**824**	**918**	**1300**	**768**	**719**	**648**	**433**	**508**
75%	395	839	980	1385	908	785	731	505	533
25%	12.7	68.9	14.0	53.7	0.0	0.0	14.0	21.4	12.7
Nourishmentforeshoreb [-]	**12.7**	**73.3**	**34.7**	**62.7**	**0.0**	**8.2**	**16.1**	**21.9**	**12.7**
75%	12.7	86.0	44.2	82.5	0.0	13.3	16.1	22.4	12.7
25%	393	2136	433	1666	0.0	0.0	433	664	393
Nourishmentforeshoreb [m^3]	**393**	**2272**	**1074**	**1943**	**0.0**	**255**	**498**	**677**	**393**
75%	393	2667	1369	2557	0.0	411	498	693	393
25%	8.2	20.6	3.6	3.0	0.0	0.0	0.0	51.1	10.4
Nourishmentdunesb [-]	**35.5**	**25.2**	**15.3**	**31.9**	**0.0**	**4.3**	**1.9**	**55.9**	**11.4**
75%	12.8	33.0	22.9	51.1	0.0	6.5	5.2	65.0	13.5
25%	254	637	112	94	0.0	0.0	0.0	1584	321
Nourishmentdunesb [m^3]	**1099**	**782**	**474**	**988**	**0.0**	**134**	**60**	**1731**	**353**
75%	396	1021	710	1584	0.0	200	160	2015	418

References

A.L. Aktinson and T.E. Baldock. *Laboratory investigation of nourishment options to mitigate sea level rise induced erosion*. Coastal Engineering, 2020.

J. Andrew, G. Coopera, and O. H. Pilkey. *Sea-level rise and shoreline retreat: time to abandon the Bruun Rule*. Global and Planetary Change 43, 2004.

L. Aragonés, J.I. Pagán, I. López, and J. C. Serra. *Depth of closure: New calculation method based on sediment data*. International Journal of Sediment Research, 2017.

T. Beuzen. *A Python package for extracting the location of dune toes on beach profile transects*. The Journal of Open Source Software, 2019.

G. T. Breard. *Evaluating Self-Or aluating Self-Organizing Map Quality Measur ganizing Map Quality Measures as Conv es as Convergence Criteria*. University of Rhode Island, 2017.

P. Bruun. *Sea-Level Rise as a Cause of Shore Erosion*. American Society of Civil Engineers Journal of the Waterways and Harbours Division, 1962.

P. Bruun. *Coastal engineering, history of. In: Beaches and Coastal Geology*. Encyclopedia of Earth Science. Springer, Boston, MA, 1982.

P.J. Cowell, P.S. Roy, and R.A. Jones. *Shoreface translation model: Computer simulation of coastal-sand-body response to sea level rise*. Elsevier Science Publischer B.V., 1992.

S. De Vries. *Lecture 4: Aeolian Coastal Landscapes*. TU Delft, 2019.

S. De Vries, M. de Schipper, M. Stive, and R. Ranasinghe. *Sediment exchange between the subaqueous and subaerial coastal zones*. Coastal Engineering Proceedings, 2010.

S. De Vries, H.N. Southgate, W. Kanning, and R. Ranasinghe. *Dune behavior and aeolian transport on decadal timescales*. Coastal Engineering, 2012.

Deltares. *Kustlijnzorg en waterveiligheid: Een verkenning van mogelijkheden om kustlijnzorg doelmatiger af te stemmen op waterveiligheid*. Deltares, 2014.

E. Diamantidou, G. Santinelli, A. Giardino, J. Stronkhorst, and S. de Vries. *An Automatic Procedure for Dune Foot Position Detection: Application to the Dutch Coast*. Journal of Coastal Research, 2020.

P. A. Estévez, J.C. Príncipe, and Pablo Zegers. *Advances in Self-Organizing Maps: 9th International Workshop*. Springer, 2012.

A. Giardino, E. Diamantidou, S. Pearson, G. Santinelli, and K. den Heijer. *A Regional Application of Bayesian Modeling for Coastal Erosion and Sand Nourishment Management*. Water, 2019.

J. Guille´n, M.J.F. Stive, and M. Capobianco. *Variability of shore and shoreline evolution*. Coastal Engineering, 2002.

B. T. Gutierrez, N. G. Plant, and E. R. Thieler. *D-separation*. Youtube, 2012.

N. G.Valiente, G. Masselink, T. Scott, D. Conley, and R. J. McCarroll. *Role of waves and tides on depth of closure and potential for headland bypassing*. Marine Geology, 2019.

L. Hamma, M. Capobiancob, H.H. Dettec, A. Lechugad, R. Spanhoff e, and M.J.F. Stive. *A summary of European experience with shore nourishment*. Coastal Engineering, 2002.

Z. Hausfather and G. Peters. *Emissions – the 'business as usual' story is misleading*. Nature, 2020.

HHNK. *Legger Zandige Kust*. HHNK, 2020.

C.L. Hinton. *Decadal morphodynamic behaviour of the Holland shoreface. PhD thesis*. Middlesex University, 2000.

IPCC, Masson-Delmotte, V., P. Zhai, H.-O. Pörtner, D. Roberts, J. Skea, P.R. Shukla, A. Pirani, W. Moufouma-Okia, R. Pidcock C. Péan, S. Connors, J.B.R. Matthews, Y. Chen, X. Zhou, E. Lonnoy M.I. Gomis, T. Maycock, M. Tignor, , and T. Waterfield (eds.). *2018: Global Warming of 1.5°C. An IPCC Special Report on the impacts of global warming of 1.5°C above pre-industrial levels and related global greenhouse gas emission pathways, in the context of strengthening the global response to the threat of climate change, sustainable development, and efforts to eradicate poverty*. In Press., 2018.

IPPC, J.A. Church, and P.U. Clark. *The Fifth Assessment Report. Chapter 13: Sea Level Change*. IPCC, 2013.

J. G. S. Keijsers, A. Poortinga, M. J. P. M. Riksen, and J. Maroulis. *Spatio-Temporal Variability in Accretion and Erosion of Coastal Foredunes in the Netherlands: Regional Climate and Local Topography*. Public Library of Science, 2014.

KNMI. *Extreme zeespiegelstijging in de 21e eeuw*. Koninklijk Nederlands Meteorologisch Instituut, 2017.

A. Kohler, M. Ohrnberger, and F. Scherbaum. *Unsupervised pattern recognition in continuous seismic wavefield records using Self-Organizing Maps*. Geophysical Journal International, 2010.

T. Kohonen. *Self-Organazing Maps (Second Edition)*. Springer, 1997.

H. Lan. *Analyzing Climate Patterns with Self-Organizing Maps (SOMs)*. Towards Data Science, 2018.

M. Larson, L. Erikson, and H.hanson. *An analytical model to predict dune erosion due to wave impact*. Coastal Engineering, 2004.

H. Li. *Which Machine Learning Algorithm Should I Use?* SAS, 2017.

I. López, L. Aragones, and Y. Villacampa. *Modelling the cross-shore profiles of sand beaches using artificial neural networks*. Marine Georesources and Geotechnology, 2018.

J. B. MacQueen. *Some methods for classification and analysis of multivariate observations*. University of California, 1967.

S.W. Marsh, R.J. Nicholls, A.Kroon, and P. Hoekstra. *Assessment of Depth of Closure on a Nourished Beach: Terschelling, The Netherlands*. Coastal Engineering, 1998.

H. Matsushita and Y. Nishio. *Batch-Learning Self-Organizing Map with Weighted Connections Avoiding False-Neighbor Effects*. IEEE World Congress on Computational Intelligence, 2010.

P. Melin, J. Cesar Monica, D. Sanchez, and O. Castillo. *Analysis of Spatial Spread Relationships of Coronavirus (COVID-19) Pandemic in the World using Self Organizing Maps*. Elsevier, 2020.

F.A.J. Minneboo. *Jaarlijkse kustmetingen: Richtlijnen voor de inwinning, bewerking, en opslag van gegevens van jaarlijkse kustmetingen*. Rijkswaterstaat, RIKZ, 1995.

P. Mitchell-Guthrie. *Looking backwards, looking forwards: SAS, data mining, and machine learning*. The SAS Data Science Blog, 2014.

A. Morang and W.A. Birkemeier. *Depth of Closure on Sandy Coasts*. Encyclopedia of Coastal Science, 2005.

M.R.Hashemia, Z. Ghadampou, and S.P. Neill. *Using an artificial neural network to model seasonal changes in beach profiles.* -, 2010.

B. Murphy. *Demystifying Buzzwords: Using Data Science and Machine Learning on Unsupervised Big Data*. Zencos, 2018.

Prof. J. Pietrzak. *An Introduction to Oceanography for Civil and Offshore Engineers*. Depth. of Hydraulic Engineering, TU Delft, 2011.

R. Ranasinghe, D. Callaghan, and M.J.F. Stive. *Estimating coastal recession due to sea level rise: beyond the Bruun rule*. Climate Change, 2012.

D. Reeve, A. Chadwick, and C. Fleming. *Coastal Engineering: Processes, Theory and Design Practice*. Spon Presss, 2002.

B.G. Reussink and M.C.J.L. Jeuken. *Dunefoot dynamics along the Dutch coast*. Earth Surface Processes and Landforms, 2002.

Rijkswaterstaat. *Kustgenese 2.0: kennis voor een veilige kust*. Rijkswaterstaat, 2020.

E. Sarkar. *Kohonen Self-Organizing Maps: A special type of Artificial Neural Network*. Towards Data Science, 2019.

M. Stewart. *The Actual Difference Between Statistics and Machine Learning*. Towards Data Science, 2019.

M.J.F. Stive, S.G.J. Aarninkhof, L. Hamm, H. Hanson, M/ Larson, K.M. Wijnberg, R.J. Nicholls, and M. Capobianco. *Shoreline evolution of the Holland coast on a decadal scale*. Earth Surface Processes and Landforms, 1999.

SuperDataScience. *EXTRA: K-means Clustering (Part 3)*. SuperDataScience Team, 2018.

H.R. Tizhoosh. *Machine Intelligence - Lecture 7 (Clustering, k-means, SOM)*. University of Waterloo, 2019.

United Nations. *The Ocean Conference*. United Nations, 2017.

C.O. Van IJzendoorn, S. de Vries, and C. Hallin. *Contribution of aeolian and marine processes to decadal coastal dune evolution along the Dutch Coast: The effect of sea level rise*. TU Delft, 2019.

H. J. Verhagen. *Coastal Protection and Dune Management in The Netherlands*. Journal of Coastal Research, 1990.

Algemene Waterschapspartij. . Waterschap, 2020.

C. Wegman and J. Leenders. *Profielverandering door suppleties en zeespiegelstijging: Kustlijn Hollands Noorderkwartier*. HKV HHNK, 2020.

X. Xia. *Self-Organizing Map for Characterizing Heterogeneous Nucleotide and Amino Acid Sequence Motifs*. University of Ottawa, 2017.

K. Zhang, B.C. Douglas, and S.P. Leatherman. *Global Warming and Coastal Erosion*. Climatic Change, 2004.