



Master of Science Thesis Geomatics

# Automatic detection of buried channel deposits using high resolution laser altimetry data (FLI-MAP)

B.M.J. Possel

December 2009





**Automatic detection of buried channel  
deposits using high resolution laser  
altimetry data (FLI-MAP)**

---

Master of Science Thesis Geomatics  
Delft University of Technology, The Netherlands

**Date:** December 16th 2009  
**By:** Boudewijn Maarten Johannes Possel  
**Professor:** Prof. dr. D.G. Simons  
**Supervisors:** Dr. R.C. Lindenbergh  
Dr. J.E.A. Storms  
**Co-Reader:** Dr. R.M. Hoogendoorn





# Preface

This report is the result of my graduation research and it is the final part of the MSc Geomatics at Delft University of Technology. The work for this research was performed at the Mathematical Geodesy and Positioning (MGP) group at the faculty of Aerospace Engineering together with the section Applied Geology at the faculty of Civil Engineering and Earth Sciences (CiTG) in combination with Fugro-Inpark B.V.

First of all I would like to thank my supervisors Roderik Lindenbergh (section Optical & Laser Remote Sensing) and Joep Storms (section Applied Geology) for the advice they have given, the discussions we have had and for the recommendations and the corrections they made to this thesis.

I would also like to thank Professor D.G. Simons (Delft University of Technology, section Acoustic Remote Sensing) and Martin Kodde (Fugro, Research & Development department) for their input. Also the help of Sabine Put (ESRI) with the online viewer and the web map service was appreciated. This also holds for Olivier Hoes (Delft University of Technology, department of watermanagement) for supplying the ArcGIS licences and the ICT department of Aerospace Engineering for all the ICT problems they had to solve for me.

There are a number of people and institutions that provided the datasets: Fugro Aerial Mapping B.V., Kim Cohen and Esther Stouthamer (Universiteit Utrecht), Harry Monis (Kadaster, sector GBKN) and Peter Jellema (DINOLoket).

I would like to thank my fellow Geodesy and Geomatics students for making my time in Delft a period that I have never regretted. Last of all, a word of thanks to my girlfriend Wendy and my parents for their continuous love and support during this master thesis.

*Boudewijn Possel  
Delft, December 2009*



# Summary

The formation of the current Rhine-Meuse delta has mainly taken place during the Holocene (the last 12,000 years). This period is characterised by avulsions that lead to sudden shifts in the location of river channels. The channel deposits left behind by abandoned rivers are often distinct sandy layers in the shallow sub-surface, called buried channel deposits. Knowledge of the location of buried channel deposits is interesting from a historical point of view and is essential for planning, constructing and maintaining structures that intersect these channels.

Traditionally, buried channel deposits are traced using labour intensive soil drillings and visual interpretation. However, some buried channel deposits can also be detected by using local elevation differences (1 - 2 m) caused by differential compaction. Automating the detection process and using Digital Elevation Models could result in faster processing and a higher detailed, more objective map. This leads to the following research question: is it possible to automatically derive a detailed map of buried channel deposits from high resolution laser altimetry data? The dataset used for this purpose is a rasterized elevation model measured by the FLI-MAP system with the same specifications as AHN-2.

First, two filtering methods are applied to remove infrastructure and other objects using existing topographical data (GBKN) and terrain characteristics (variability and density). If these objects are not removed they cause errors in the final result. In the detection procedure four structural attributes are calculated using the elevation dataset: slope, curvature, relative elevation (TPI) and smoothed TPI. This is done to use as much of the available information in the detection procedure. With these attributes a multi-band image is formed. Classification is then performed using a Maximum Likelihood classifier where each point is assigned to a predefined class. A majority filter is applied on the result and empty areas are interpolated using a conditional dilation to get a full coverage of the area. Processing speed is increased by downsampling datapatches that are used for computational intensive algorithms. Additional automation of the detection process is performed by developing a method to resolve edge effects on tiled datasets.

Validation is performed by comparing the detection result to two independent datasets. A palaeogeographic map [5] is used to make a numerical and visual comparison. An analysis using shallow drilling measurements is performed to determine the depth of the detected sand layers.

The results indicate that buried channel deposits (between 4,620 and 1,700 years old) can be mapped roughly in a fully automatic way. The probability that these channels are within a certain depth can be calculated. Validation shows that this new method is strongly hampered due to human intervention in the landscape. Further research is recommended to improve the filtering and detection method.





# Samenvatting

De vorming van de huidige Rijn-Maas-delta heeft voornamelijk plaatsgevonden gedurende het Holoceen (de laatste 12.000 jaar). Deze periode wordt gekenmerkt door avulsies die leiden tot plotselinge verschuivingen van een rivierloop. Achtergelaten rivierenbeddingen zijn vaak als zanderige lagen terug te vinden in de ondiepe ondergrond, zogenaamde ondiepe zandbanen. Kennis over de locatie van deze zandbanen is interessant vanuit een historisch oogpunt en is essentieel voor de planning, bouw en het onderhoud van infrastructuur die kruist met deze banen.

Traditioneel worden zandbanen opgespoord met behulp van arbeidsintensieve grondboringen en visuele interpretatie. Een gedeelte van de zandbanen kan worden opgespoord door gebruik te maken van lokale hoogte verschillen (1 - 2 m), veroorzaakt door differentiële compactie. Automatiseren van het proces van opsporing en het gebruik van Digitale Hoogte Modellen zou kunnen leiden tot een snellere verwerking en een meer gedetailleerde, objectieve kaart. Dit resulteert in de volgende onderzoeksvraag: is het mogelijk om automatisch een gedetailleerde kaart van ondiepe zandbanen af te leiden uit hoge resolutie laser altimetrie data? De dataset gebruikt voor dit doel is een gerasterde hoogte dataset gemeten door het FLI-MAP-systeem, met dezelfde specificaties als AHN-2.

Eerst worden twee filters toegepast om infrastructuur en andere objecten te verwijderen, dit gebeurt met behulp van bestaande topografische gegevens (GBKN) en de karakteristieken van het terrein (variabiliteit en dichtheid). Als deze objecten niet worden verwijderd veroorzaakt dit fouten in de eindresultaten. In de detectie procedure worden vier structurele attributen berekend uit de hoogte dataset: helling, kromming, relatieve hoogte (TPI) en gemiddelde TPI. Dit wordt gedaan om zoveel mogelijk van de beschikbare informatie te gebruiken in de detectie procedure. Uit deze attributen wordt een multi-spectraal beeld samengesteld. De classificering wordt hierna uitgevoerd met behulp van een "Maximum Likelihood" schatter waarbij elk punt wordt toegewezen aan een vooraf bepaalde klasse. Hierna wordt een "majority filter" toegepast en worden lege gebieden geïnterpoleerd met behulp van een voorwaardelijke dilatatie. De verwerkingssnelheid wordt verhoogd door het reduceren van de hoeveelheid datapunten in gebieden die worden gebruikt voor berekenings intensieve algoritmen. Tevens is een methode ontwikkeld die randeffecten opheft die ontstaan bij het verwerken van getegelde datasets.

Validatie wordt uitgevoerd door de resultaten met twee onafhankelijke datasets te vergelijken. Een paleogeografische kaart [5] wordt gebruikt voor een numerieke en visuele vergelijking. Ook wordt een analyse uitgevoerd met behulp van ondiepe boringen om de diepte van de gedetecteerde zandbanen te bepalen.

De resultaten laten zien dat ondiepe zandbanen (tussen 4.620 en 1.700 jaar oud) globaal kunnen worden gedetecteerd op een automatische manier. De kans dat zandbanen binnen een bepaalde diepte liggen kan worden berekend. Uit validatie blijkt dat de detectiemethode sterk wordt belemmerd als gevolg van menselijk ingrijpen in het landschap. Verder onderzoek wordt aanbevolen met betrekking tot het filteren en de detectiemethode.



# Abbreviations

AHN-1	Actueel Hoogtebestand Nederland 1 (Actual Height model of the Netherlands 1)
AHN-2	Actueel Hoogtebestand Nederland 2 (Actual Height model of the Netherlands 2)
ALS	Airborne Laser Scanning
API	Application Programming Interface
CPT	Cone Penetration Test
DEM	Digital Elevation Model
DINO	Data en Informatie van de Nederlandse Ondergrond (Data and Information of the Subsurface of The Netherlands)
DSM	Digital Surface Model
FLI-MAP	Fast Laser Imaging - Mapping And Profiling
GB	Gigabyte
GBKN	Grootschalige BasisKaart Nederland (The Large Scale Standard Map of The Netherlands)
GIS	Geographic Information System
GDAL	Geospatial Data Abstraction Library
GPS	Global Positioning System
IMU	Inertial Measurement Unit
Laser	Light Amplification by Stimulated Emission of Radiation
NAP	Normaal Amsterdams Peil (Amsterdam Ordnance Datum)
NP	Navigation Processor
Pixel	Picture Element
RD	Rijksdriehoekstelsel (Dutch National Coordinate System)
Radar	Radio Detection And Ranging
TPI	Topographic Position Index



# List of Symbols

$\mathbf{A}$	Model matrix
$A$	Shape
$B$	Structuring element
$C$	Conditional shape
$\partial$	Derivative
$c_i$	Plane parameter
$\hat{c}_i$	Best-fitting plane parameter
$d_i$	Curvature parameters
$\hat{d}_i$	Best-fitting curvature parameters
$d_{ii}$	Diagonal matrix elements
$d_{i+}$	Sum of elements in a matrix row
$d_{j+}$	Sum of elements in a matrix column
$\Delta$	Difference
$F()$	Cumulative distribution function
$H$	Mean curvature
$i$	Distance in pixels
$k_1$	Maximal curvature
$k_2$	Minimal curvature
$\hat{\kappa}$	Kappa coefficient (statistical measure)
$L$	Window size
$M$	Distance from centre pixel to the edge of the window
$N$	Selection of points
$N_{tot}$	Total number of points
$N_i$	Adjacency
$O$	Computational complexity
$P()$	Probability of an event
$p$	Horizontal size of data tile
$q$	Vertical size of data tile
$r$	Radius of inner circle
$R$	Radius of outer circle
$s$	Slope of fitted plane
$t$	Tile number
$\tau_{tpi}$	TPI-value
$\overline{\tau_{tpi}}$	Average TPI-value



$x$	Vector with unknown parameters
$\hat{x}$	Vector with Best Linear Unbiased Estimation of the parameters
$x_i$	X-coordinate
$\bar{x}$	Average x-coordinate
$y_i$	Y-coordinate
$\bar{y}$	Average y-coordinate
$z_i$	Height (relative to NAP)
$z_{tpi}$	Height of the centre pixel used for TPI
$z_x$	Partial derivative of z
$z_y$	Partial derivative of z
$z_{xx}$	Partial derivative of z
$z_{xy}$	Partial derivative of z
$z_{yy}$	Partial derivative of z

# Contents

<b>Preface</b>	<b>v</b>
<b>Summary</b>	<b>vii</b>
<b>Samenvatting</b>	<b>ix</b>
<b>Abbreviations</b>	<b>xi</b>
<b>List of Symbols</b>	<b>xiii</b>
<b>1 Introduction</b>	<b>1</b>
1.1 Background . . . . .	1
1.2 Research objectives . . . . .	3
1.3 Methodology . . . . .	4
1.4 Structure . . . . .	5
<b>2 Fluvial Palaeogeography</b>	<b>7</b>
2.1 River channels . . . . .	7
2.1.1 River Morphology . . . . .	8
2.1.2 Avulsion . . . . .	9
2.2 Buried channel deposits . . . . .	10
2.2.1 Topographic Inversion . . . . .	10
2.2.2 Characteristics of buried channel deposits . . . . .	11
2.3 Location . . . . .	12
2.3.1 History . . . . .	13
<b>3 Data</b>	<b>15</b>

3.1	Principle of Airborne Laser Scanning . . . . .	16
3.1.1	FLI-MAP . . . . .	17
3.1.2	FLI-MAP Data Processing . . . . .	17
3.1.3	Additional FLI-MAP Data . . . . .	19
3.2	GBKN . . . . .	20
3.3	Digital palaeogeographic map . . . . .	20
3.4	DINOLoket . . . . .	21
<b>4</b>	<b>Filtering procedure</b>	<b>23</b>
4.1	Filtering using GBKN . . . . .	23
4.1.1	Snapping . . . . .	23
4.1.2	Self Intersection . . . . .	24
4.1.3	Area building . . . . .	24
4.1.4	Buffering . . . . .	24
4.1.5	Further processing . . . . .	25
4.2	Variability filter . . . . .	27
4.2.1	Variability . . . . .	28
4.3	Density filter . . . . .	30
4.4	Results . . . . .	30
4.4.1	Remarks . . . . .	31
4.4.2	Discarded filters . . . . .	33
<b>5</b>	<b>Detection</b>	<b>35</b>
5.1	Structural Attributes . . . . .	35
5.1.1	Attribute 1 - Slope . . . . .	36
5.1.2	Attribute 2 - Curvature . . . . .	38
5.1.3	Attribute 3 - TPI (Topographic Position Index) . . . . .	40
5.1.4	Attribute 4 - Smoothed TPI . . . . .	41
5.2	Classification . . . . .	41
5.2.1	Multispectral image classification . . . . .	42
5.3	Downsampling . . . . .	44
5.3.1	Moving window . . . . .	44
5.3.2	Point downsampling for plane fitting . . . . .	45
5.3.3	Downsampling optimization . . . . .	46

---

5.3.4	Downsampling for other attributes . . . . .	47
5.4	Resolving Edge Effects . . . . .	48
5.4.1	Moving window operations on tiled datasets . . . . .	48
5.4.2	Topology . . . . .	50
<b>6</b>	<b>Results and Validation</b>	<b>51</b>
6.1	Structural attributes . . . . .	51
6.2	Results of detection . . . . .	54
6.3	Filtering and Interpolation . . . . .	55
6.3.1	Filter . . . . .	56
6.3.2	Conditional dilation . . . . .	56
6.4	Validation . . . . .	59
6.4.1	Comparison with palaeogeographic map . . . . .	59
6.4.2	Comparison with DINO Data . . . . .	63
6.5	Final product . . . . .	66
6.5.1	Visualisation . . . . .	66
6.5.2	Toolbox . . . . .	67
<b>7</b>	<b>Conclusions and Recommendations</b>	<b>69</b>
7.1	Conclusions . . . . .	69
7.1.1	General . . . . .	69
7.1.2	Filtering . . . . .	70
7.1.3	Detection . . . . .	70
7.1.4	Validation . . . . .	70
7.1.5	Final product . . . . .	71
7.2	Recommendations . . . . .	71
	<b>Bibliography</b>	<b>73</b>
	<b>Appendix A - Workflow Diagram</b>	<b>77</b>
	<b>Appendix B - Software</b>	<b>79</b>
	<b>Appendix C - Borehole description example</b>	<b>81</b>
	<b>Appendix D - Processing time</b>	<b>83</b>

---

<b>Appendix E - Discarded Processing techniques</b>	<b>85</b>
<b>Appendix F - Elevation dataset of the Alblasserwaard</b>	<b>87</b>
<b>Appendix G - Field measurements</b>	<b>89</b>

# Chapter 1

## Introduction

### 1.1 Background

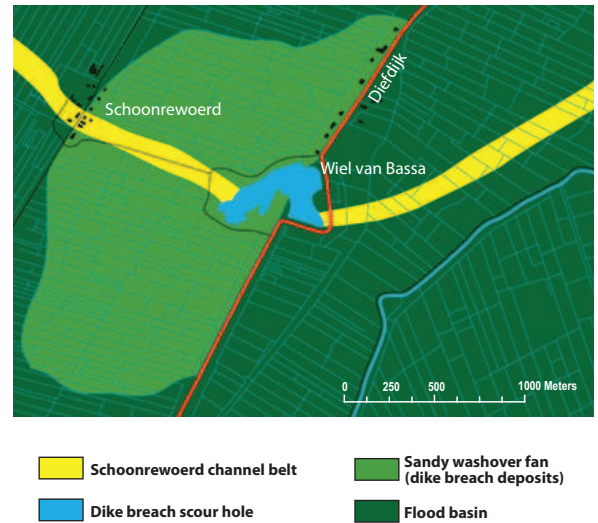
Knowledge of the complex subsurface of the Netherlands is very important. Information about the subsurface is used in many different applications: for instance in planning and performing construction projects like highways, tunnels, railroads and new housing developments. Surface area in the Netherlands is limited and used intensively, this has led to higher usage of the space underground. All of these reasons makes us want to know more about the structure and processes that occur below our feet.

If we want to understand more about the subsurface we have to go back to the period that it was formed. A large part of The Netherlands as we know it today, was formed during the Holocene (the period from 12,000 years ago until present). Water flowing in rivers carried materials like clay and sand. In the past these rivers often changed their location, abandoning their old channels and forming new ones. The deposits left behind are often distinct sandy layers in the shallow sub-surface, called buried channel deposits. The location of these channels deposits are of interest for of a number of reasons:

- First of all, it is interesting from a cultural and historical perspective. The location of these channels deposits can teach us about the history of the Netherlands, help us understand what processes took place and how this influenced the landscape and what impact this had on our ancestors.
- The location of buried channel deposits is relevant for the construction and maintenance of roads and other structures that cross these channels. The difference in compaction of the underlying sediments causes height differences at the surface. This effect can cause increasingly bumpy roads as time progresses. Examples of this effect in the Netherlands are the A15 near Ridderkerk [23] and the Vogelweg in southern Flevoland [17] & [21], seen in Figure 1.1. Differential compaction can also influence other structures like railroads, sewers and even buildings.



**Figure 1.1:** Natural speed bumps have appeared in the Vogelweg in Southern Flevoland due to an ancient buried channel deposits made by the river Eem [21].



**Figure 1.2:** Dike burst of the Dierdijk in the east of the Alblasserwaard, caused by seepage that undermined the dike. Modified from [3].

- The intersections of dikes with buried channel deposits are of particular interest, because at these locations dikes are more vulnerable to seepage or piping. This is caused by the fact that water flows more easily through sand, where it will gradually carry away the sand particles causing channels that undermine the dike. It is thought that up to 80% of dike failures have been caused by this effect at the location of buried channel deposits. Remnants of these failures can still be seen today in the landscapes, an example is the 'Wiel of Bassa' [42] seen in Figure 1.2. Information about the location of these weak spots can be used for additional inspection, maintenance or strengthening of existing dikes. The location of buried channel deposits can also be used for the optimal planning of new dikes [27].
- The water permeability properties of buried channel deposits cause them to act as pipelines for groundwater, this is called an aquifer. Care must be taken when excavation works are initiated in these areas, because water could be transported to the location where the buried channel deposits are uncovered, resulting in unwanted wells. Water contained in the sand layer, the surrounding area or in deeper lying groundwater systems is then extracted. This could influence the groundwater level of the surrounding area, locally resulting in subsidence and drought. In some locations this effect could cause salt water intrusion through buried channel deposits [29], leading to ecological damage.
- Finally, the locations of buried channel deposits can be used for subsurface Geotechnical investigations. Geotechnical investigations are performed to obtain information on the physical properties of soil and rock around a site. Using the locations of buried channel deposits a more complete and detailed overview of a location can be made, this could reduce the costs or improve the planning of such investigations.



Because of the reasons above it would be preferable to know the exact locations of all the buried channel deposits. This has partially been accomplished by Utrecht University, where a detailed palaeogeographic map of the Rhine-Meuse delta (including the approximate locations of buried river channel deposits) has been constructed [5]. The map is based on more than 25 years of field research using over 200,000 manual boreholes. This is a time-consuming process that requires a large amount of manual labour in the field. For this reason not the whole Dutch river delta has been covered in the same amount of detail. The map also still contains errors caused for instance by differences in interpretation.

Buried channel deposits can be detected nowadays by making use of relative height differences in the landscape. These height differences have been measured using high-resolution airborne laser scanning, creating a Digital Elevation Model (DEM) of the area. The process of verifying buried channel deposit maps and enhancing their position accuracy could significantly be supported by the application of a DEM [6]. A DEM of the Netherlands (the AHN-1) has already been used to extract information about the location of buried channel deposits, mostly as a verification technique. This is done based on visual interpretation by a human operator and the information is used to select drilling locations and areas of interest. However, GIS technology could make it possible to automate the process of detecting buried channel deposits. Compared to the AHN-1 a more accurate DEM has been measured using FLI-MAP (Fast Laser Imaging - Mapping and Profiling). This measurement method will be used to measure the new AHN-2. For this research project it was possible to use such a high resolution DEM of an area in the Netherlands.

## 1.2 Research objectives

Until now no research has been done on automatically deriving buried channel deposit locations from a DEM. Doing this process automatically could result in a faster, more detailed and more objective derivation compared to using visual interpretation. Current maps could be compared and updated using the detection results, giving indications where additional soil measurement should be performed. Also uncharted areas (in the Netherlands or abroad) could be classified in a fast and efficient way. In cases where automatic detection is not possible and operators still need to interpret the data, specific processing techniques applied in the research could be used to assist in this interpretation.

This leads us to the main research question of this Master thesis research:

*"Is it possible to automatically derive a detailed map of buried channel deposits using high resolution laser altimetry data?"*

The main research question is subdivided into the following objectives:

1. **Filtering the FLI-MAP dataset.** Various objects like roads, trenches and buildings need to be removed from the FLI-MAP dataset because they do not belong to buried channel deposits. If unaddressed these objects will complicate or hinder the detection process. We need to determine the characteristics of these objects and choose which filtering methods subsequently need to be applied. In addition, it will be investigated how the GBKN (the Large Scale Standard Map of The Netherlands) needs to be adapted before it can be used in the filtering process.

2. **Automated channel detection.** Algorithms based on image-processing, GIS and Remote Sensing techniques will be developed and adapted to detect buried channel deposits. Questions that will be addressed are: Which kind of algorithm is suited for the detection? How can this process be automated?
3. **Verification of the detection result.** Results will be verified using independent datasets containing information about the subsurface of the area. It will be investigated if these datasets can be used to make a description of the quality of the detection result. The differences encountered between the validation datasets and the detection result will be analyzed. Using the verification dataset questions will be answered are: Until what age are buried channel deposits still detectable? What is the depth of the detected buried channel deposits?
4. **Creation of a geo-referenced channel map product.** Using data from the previous steps we want to derive a geo-referenced map of the detected buried channel deposits, possibly with an indication of the detection (un)certainty. It will be investigated how this map product can be visualized and distributed/stored.

### 1.3 Methodology

The research is performed in a case study consisting of the area of the Alblasserwaard (the Netherlands), seen in Figure 1.3. This area is located in the centre part of the Rhine-Meuse fluvial system, where the location and age of most of the larger buried channel deposits is known. To answer the research questions, various processing techniques from the field of GIS and Remote Sensing are combined: for filtering the FLI-MAP dataset suitable filtering methods have to be implemented. For the detection of buried channel deposits existing algorithms for surface extraction and classification will be used and extended. New methods have to be created for automating the whole detection process. The validation process will be done by comparing the results to independent datasets (palaeogeographic map, shallow depth drillings and Cone Penetration Tests).

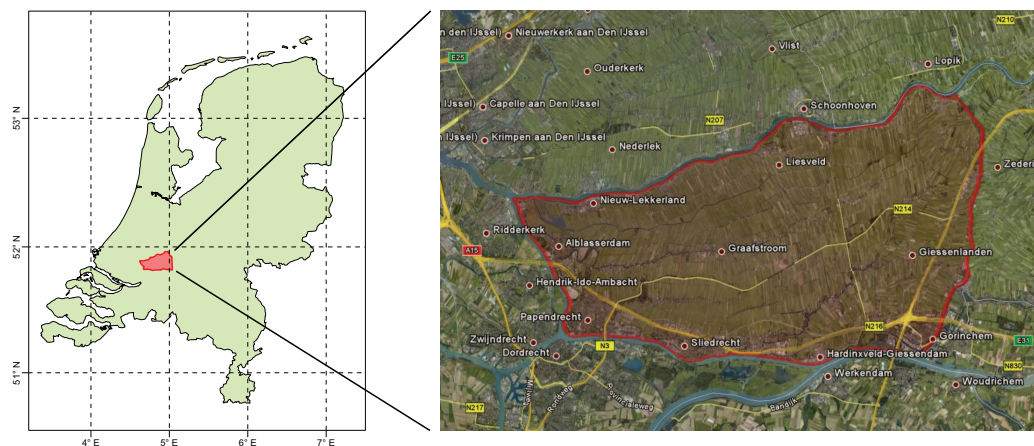


Figure 1.3: The area of the Alblasserwaard.

## 1.4 Structure

In this report the following chapters are defined:

Chapter 2 will give background information about the different geo-hydrological processes that have created the present day landscape in the case study area.

In chapter 3 background information is given about the area where this research is performed. The measurement method of airborne laser altimetry is explained, more specifically the FLI-MAP system. The procedures are described that lead to the final dataset that was used in this research. Also various other relevant datasets used in this research are discussed.

Chapter 4 describes various filtering methods which are applied to the FLI-MAP data to remove objects like roads, trenches and buildings which do not belong to buried channel deposits. If these objects are not removed they influence the detection procedure and cause errors in the final result.

Chapter 5 describes a method for the detection of buried channel deposits from the filtered FLI-MAP data. A range of algorithms and their characteristics are explained, followed by the classification of the area. Finally some additional processing steps are explained that were developed to make the procedure work automatically.

The results of the detection method are presented in chapter 6. Additional improvements are made to the detection result. For the validation process various datasets are used to determine what the quality of the result is and what the main differences are.

Finally, chapter 7 formulates conclusions and gives recommendations for future research.



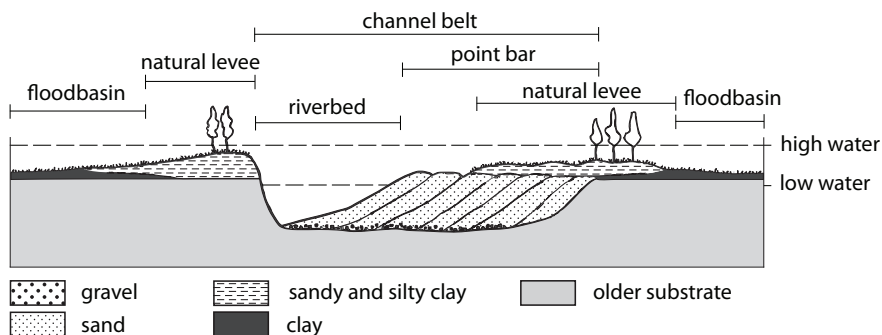
## Chapter 2

# Fluvial Palaeogeography

Palaeogeography is the study of the Earth and its lands, features, inhabitants, and phenomena. The term 'fluvial' is used to refer to the processes associated with rivers and streams and their deposits. For this research we are interested in the fluvial part of palaeogeography, so this chapter provides background on rivers, their morphology and the formation of their deposits. Also background information is given on the area and history of the location used for this research.

### 2.1 River channels

A river is a natural stream of freshwater, flowing toward an ocean, a lake or another stream. Water is collected from precipitation, surface runoff, groundwater and from reservoirs (including natural reservoirs like glaciers). Some of the terminology used when talking about rivers is shown in Figure 2.1.



**Figure 2.1:** Schematic cross section at a bend of a river (adapted from [4], after [30]).

The riverbed is the bottom of a river where water flows during low water. On both sides, running parallel with the river, lie natural levees. They are formed during flooding when coarser sediments are deposited at the edge of the channel. Over time, the levees are built up above the rest of

the surrounding area. The land behind the levees is called the floodbasin, it is inundated during high water. Further from the river the water flows slower so the deposited sediments become finer (smaller), immediately next to the river the coarser materials like fine sands and silts are deposited. Further away (usually several hundreds of meters) clay is deposited, after that no suspended material can be deposited, so peat will form. Boundaries between the different deposited materials is very gradual.

When a river meanders, point bars are formed at the inner part of a stream bend, these are composed of sandy sediment. Point bars are crescent-shaped (the same form as a partially lit moon) and have an elevation very close to the low water level. The term channel belt is applied to the whole body of sediment (mostly sand with sometimes gravel) deposited by the channel, natural levees and deposits in the floodbasin are not included. Channel belts are broader than the channel that deposited them.

### 2.1.1 River Morphology

Rivers do not only transport water, but also sediments like gravel, sand, silt and clay. As the water flows over the riverbed it removes sediment (erosion) from its original location (usually the outer banks) and transports it to a new location where it is deposited. The sediment transport rate depends on the availability of sediment and on the discharge of the river. Due to erosion, transport and deposition of sediments rivers will adopt certain patterns, depending on the regional topography and type of sediment. Rivers channels can be subdivided into four different patterns; straight, meandering, braided and anastomosing [24], [32], see Figure 2.2. Characteristics of the four different channel patterns are:

- **Straight:** Straight rivers have only one single-channel and little or no sand is deposited on the river banks. The river does not literally have to be straight, some sinuosity may occur.
- **Meandering:** These rivers only have one single channel but continuously shift their location due to the process of erosion and deposition of sediments.
- **Braided:** Braided rivers have a single channel, but multiple threading rivers. These threads migrate and are generally wide, shallow and enclose sandy bars which can be flooded during high water.
- **Anastomosing:** These rivers are composed of several interconnected channels which generally do not migrate.

Transitions can occur between the various types of patterns at different locations along the river, but this can also change during time. Channel patterns are based on the amount of water discharge, slope and sediment (particle size and amount). Due to the different properties, each of the channel types result in the formation of different subsurfaces, also seen in Figure 2.2. In the Alblasserwaard the occurring channel patterns changed during time. More than 10,000 years ago there was a braided pattern, from 10,000 to 7,500 years ago this changed into a meandering pattern. It then changed to a straight pattern which has remained until the present [5].

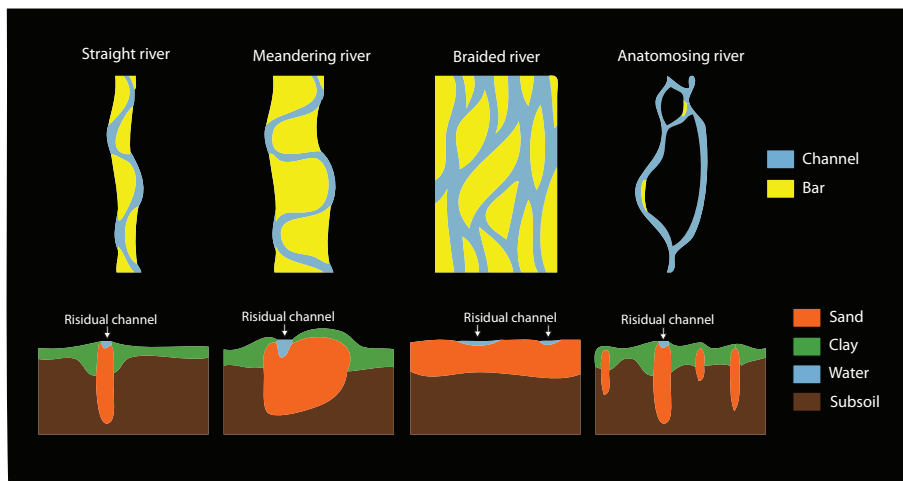


Figure 2.2: Different channel types including the formed subsurface, adapted from [39].

### 2.1.2 Avulsion

Avulsion is the abandonment of a part or the whole of a channel belt by a stream in favor of a new course [1]. This effect occurs in all kinds of fluvial environments; in straight, meandering, braided and anastomosing rivers.

If an avulsion occurs during excess discharge and floodwater spills over the natural levee, a small channel is formed called a crevasse channel. It transports sands from the riverbed into the adjacent floodbasin, initially forming a new sheet of material (splay) over the floodbasin deposits. In most cases the crevasse channel is then abandoned, leaving the old river channel intact. But sometimes a new stable channel is formed and the old river channel is (partially) abandoned. The newly formed channel will usually take a different path, because it is unlikely that it will flow back to the higher elevated levees. An example of an avulsion with accompanying crevasse splay can be seen in Figure 2.3.

From 10,000 to 3,700 years ago the Rhine-Meuse delta was characterized by large-scale crevassing and splay forming. The consecutive avulsions resulted in a complex subsurface of alternating sandy channel areas and clay/peat rich floodplain deposits. Many remnants of buried channel deposits, crevasses and splays are present in the subsurface of the Rhine-Meuse delta and have been mapped [5]. However in the Alblasserwaard few crevasse splays are indicated on the map. This does not mean that crevasse splays are not present in the Alblasserwaard but they have not been mapped yet due to, for example, their relative small width.

Besides avulsions, another process that plays a role in the river morphology is known as a meander cutoff, this process can be seen in Figure 2.4. A meander is the large bend in a meandering river, formed by erosion of the outer banks and sedimentation of the inner banks. The process of erosion and sedimentation in a meander continues until the channel meets itself and the old meander bend is abandoned.





**Figure 2.3:** *An example of the formation of a crevasse splay in the Columbia River, British Columbia, Canada [picture by H.J.A. Berendsen].*

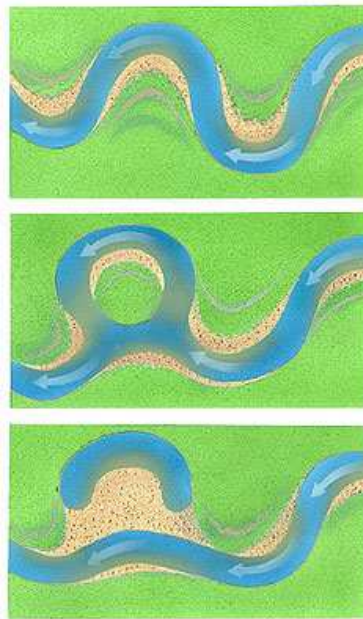
## 2.2 Buried channel deposits

As described in the previous paragraph, rivers can change their path, abandoning their old channels and forming new ones. The channels belt deposits left behind are often distinct sandy layers in the shallow sub-surface, called buried channel deposits. A total of 206 Holocene channel belts have been recorded in the Rhine-Meuse delta. Due to meandering of the river the buried channel deposits can be quite a bit broader compared to the river that deposited material, the difference between the riverbed and the channel belt can be seen in Figure 2.1. Buried channel deposits are bordered on both sides by floodplain deposits, containing clay, silt and peat rich material. The transition between the buried channel deposits and the floodplain deposits is abrupt. An example of buried channel deposits and the surrounding deposits can be seen in (Figure 2.5). A borehole description at a location of buried channel deposits can be seen in Appendix C, the sandy layers are at the greatest depth with peat and clay layers on top.

The different buried channel deposits in the Alblasserwaard were each formed during different stages in the Holocene. This causes buried channel deposits to each have their own age. Because river deposits are layered on top of each other, it can happen that buried channel deposits with different ages intersect with each other at different depths.

### 2.2.1 Topographic Inversion

Presently in the Alblasserwaard, buried channel deposits are recognizable in the landscape due to a process called topographic inversion. This occurs when peat, silt and organic material on both sides of the buried channel deposits decay or compact at a higher rate than the sand in the channel itself. At the surface this causes an area with a higher elevation at the locations of buried channel deposits. Topographic inversion can be observed in Digital Elevation Models (DEM), for instance in Figure 2.6. However, the buried channel deposits have a much higher density compared to the layers of clay, silt or peat that surround it. The weight of the buried channel deposits causes compaction of the underlying layers, causing partial compensation of the process of topographic inversion.



**Figure 2.4:** *Process of a meander cutoff [9].*

When, due to an avulsion or a meander cut off, water no longer flows through a channel it is called an abandoned channel. Abandoned channels fill up with clay and peat. This material also compacts during time, far more than the levee and pointbar deposits. So again due to differential compaction the abandoned channel can sometimes be seen in a DEM, but now as a slight depression inside of the elevated area caused by the buried channel deposits. This effect can be seen at certain locations in the Alblasserwaard.

### 2.2.2 Characteristics of buried channel deposits

To use the characteristics of buried channel deposits in the detection process a list of properties has been made. In an ideal situation, buried channel deposits would have the following characteristics:

1. If differential compaction occurs then buried channel deposits cause the surface to have a **higher elevation** compared to the surrounding area. Various factors lead to the amount of surface elevation difference: the thickness and amount of compaction of deposited materials, the amount of oxidation of organic materials and the amount of compaction of lower-lying layers due to the weight of the buried channel deposits.
2. Due to the low surface gradient of the Rhine-Meuse delta, the rivers deposited material very evenly. This is why individual buried channel deposits lie at a relatively constant depth, which in turn causes a **constant surface elevation**.
3. The buried channel deposits have well defined borders, but because the layers don't abruptly end vertically, there will be a gradual, but defined **transition zone** at the surface.

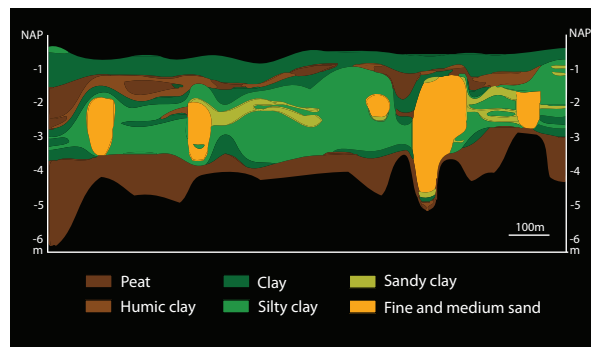


Figure 2.5: Cross section of the different layers in the subsoil, located at Lopik (adapted from [35]).

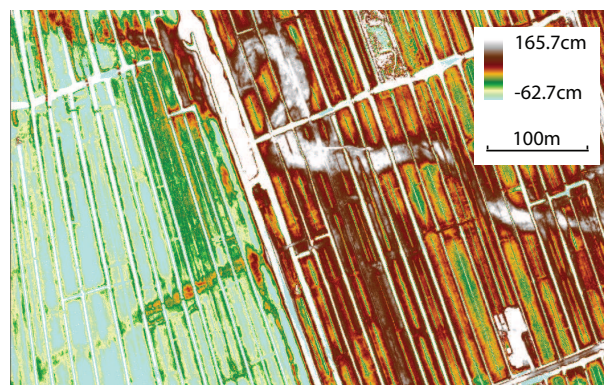


Figure 2.6: Former river channel visible in a DEM.

4. Buried channel deposits occur at **various scales**. The sizes of channels belts vary, the main channels belts have an average cross-section of 100 - 300 m, side channels can be 20 - 100 m across.
5. **Connectivity**, fluvial channels are continuous (will not suddenly end) or are connected to each other. Tidal creeks or crevasse splays are mostly connected on a single side only.
6. **Anisotropy** is the property of being directionally dependent, so there is no homogeneity in all directions. Anisotropy in surface elevation by buried channel deposits is caused by the rivers flowing in a certain direction. For the Alblasterwaard this is mostly East-West, but locally the direction is variable due to meandering and the presence of perpendicular side channels and smaller tidal creeks.

## 2.3 Location

The area used in the case study for this thesis project is the Alblasterwaard. This is a region located in the south of the province of Zuid-Holland, the Netherlands (see Figure 1.3). To the west are the cities of Rotterdam and Dordrecht, Utrecht is located a little further to the north-east. It currently has

about 170,000 inhabitants and its surface amounts to approximately 350 km<sup>2</sup>. The Alblasserwaard can roughly be divided into two areas: the urbanized area in the west together with the southern part along the river, and the mainly rural area largely covered by grassland located in the centre, northern and eastern part of the Alblasserwaard. The largest towns in the Alblasserwaard are Gorinchem and Papendrecht.

The Alblasserwaard lies below sea level, on average approximately -1.20 m below NAP (Amsterdam Ordinance Datum). Currently about 20 different waterlevels are maintained throughout the polder, ranging from -0.80 m to -2.00 m [14]. The Alblasserwaard is bordered by rivers on all sides; to the north by the river Lek, in the south by the river Merwede, in the west by the river Noord, in the east by the old Zederik, the Merwedekanaal and the Linge. Because the area was entirely surrounded by water it is known as a 'waard'. The rest of the name of the Alblasserwaard was derived from the 'Alblas' effluent river.

### 2.3.1 History

The (sub)surface of the Alblasserwaard and surrounding area has been formed by many different processes during different times. The area as we see it today is an accumulation of all these different processes. Two relevant periods in history of the Alblasserwaard that will be described in the next paragraphs are:

- **Holocene until Early Middle Ages (12,000 years ago until 1,000 years ago)**  
The period from 12,000 years ago until 1000 years ago in the Alblasserwaard is characterized by natural processes. In this period buried channel deposits were formed, so information about their formation could lead to an improvement in the detection algorithms.
- **Early Middle Ages until present (1,000 years ago until present)**  
This period describes the more recent history of the Alblasserwaard (1,000 years ago until present), it is characterized by human influence. It accounts for many of the small-scale surface features that can be seen today, knowledge about this period can possibly be used to filter the height dataset.

The very early history of the Alblasserwaard (Pleistocene and before, more than 12,000 years ago) is not covered. Knowledge about this period is important to understand how the Netherlands was created due to the formation of for instance the underlying layers. But we will use this information as a starting point, and not go into details about how the underlying layers were formed.

The subsurface of the Alblasserwaard and the surrounding area were formed during the Pleistocene (1.8 million to 12,000 years ago). It consists of sands, formed during the last ice-age in the Pleistocene (21,000 years ago).

#### **Holocene until Early Middle Ages**

Ice-age Pleistocene sands (116,000 until 11,500 years ago) from the bottom of the North Sea were transported and deposited, resulting in the formation of barrier beaches and coastal dunes in the

western part of the Netherlands. Because of this the sea had less influence on the back-barrier area, in combination with the relative sea level rise it led to groundwater level rise causing the formation of a series of fresh-water marches. In the marches partially decayed vegetation matter accumulated, causing the formation of peat. In the Alblasserwaard the peat forming process was influenced by the rivers, regular inundations caused clay to be deposited with the peat, forming very fertile grounds that covered the low lying Pleistocene deposits. In the course of time many layers were laid down on top of each other. The thickness of the peat and clay deposits varies from about 1m near the German border to about 25 m near the Dutch coast [5]. The vegetation present in the Alblasserwaard was forest, causing the formation of wood peat. This type of peat is not suited for use as fuel. Together with the large amount of clay mixed with the peat, this is the reason why there has never been large-scale peat excavation in the Alblasserwaard [43].

Tidal differences of up to 4 m at the coast [5] resulted in the formation of fresh water tidal creeks in the Alblasserwaard. This area was probably comparable to the present day area of the Biesbosch. Due to the swampy nature of the area and the frequent floods, large parts of the Alblasserwaard remained uninhabited. The first settlers moved to highest lying areas; old riverdunes (in Dutch: "donken"), levees and crevasse deposits. Most of the Alblasserwaard remained uninhabited, however large scale habitation occurred during the Roman period. A depopulation followed after the year 250 AD.

### **Early Middle Ages until present**

The Alblasserwaard became permanently inhabited from the year 1,000 onwards, when land reclamation projects started. The reclamation was done according to the cope-system, which is in fact an agreement between a land-lord and a group of workers. The standard size of a parcel was set at 1,250 m by approximately 112 m, but the length of the parcels was mainly influenced by the presence of rivers or creeks. Drainage ditches were dug around the parcels to remove excess water. The elongated shapes of these parcels can still be seen today.

The reclamation process first started from the sides of the polder where the main rivers flowed. Reclamation then continued inwards, using the existing streams that ran through the centre of the Alblasserwaard from east to west. Parcels were created perpendicular to those canals and parallel to each other. Houses and farms were built along side of the canals causing ribbon development (in Dutch: "lintbebouwing"). The reclamation of the Alblasserwaard was probably completed in the fourth quarter of the 13<sup>th</sup> century.

Besides the reclamation there were also other human influences: rivers were embanked and meanders were cut off. This stopped the large scale sedimentation of the area, however the polders were still inundated many times due to dike failures. Water extraction caused the peat to subside, which in turn caused the central part of the Alblasserwaard to become even more prone to flooding. Water management became very important, especially in combination with rising waterlevels of the rivers. Excess water was removed from the low lying areas using a system of waterways and pumps. This was formerly done by using the windmills on Kinderdijk.

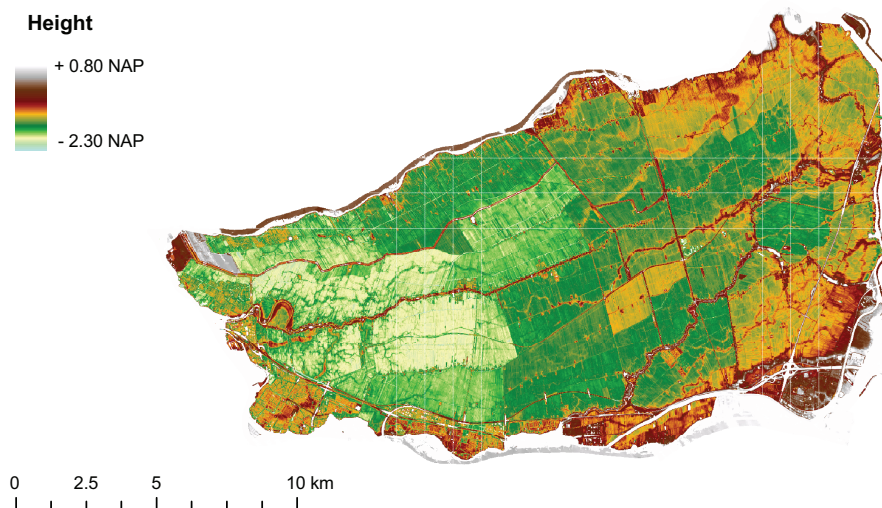
In the 1980's large scale agricultural consolidation was performed, farmers traded their land so scattered plots of farmland could be formed into larger and more efficient parcels. Also during this time leveling/equalization was applied to certain parcels, causing the loss of relief in the landscape.

## Chapter 3

# Data

This chapter gives information on the main elevation dataset and the various other datasets used in this research. Also the principle of Airborne Laser Altimetry (ALS) is explained and how this is applied in the FLI-MAP system.

The main dataset used in this research is FLI-MAP elevation data, measured by Fugro Aerial Mapping B.V. The owner of the dataset is the Waterboard "Rivierenland". Both the Waterboard and Fugro are interested in the location of buried channel deposits. An overview of the entire dataset can be seen in Figure 3.1. The data were acquired during three days in August 2007 with a minimum point density of 8 - 10 point per m<sup>2</sup>. The pixels in the original rasterized dataset have a size of 0.5×0.5 m, but have been resampled to 2×2 m for further processing.

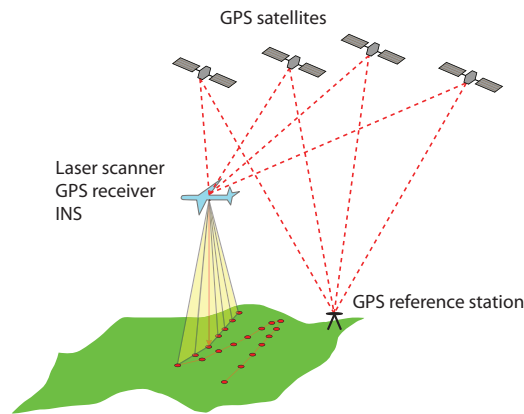


**Figure 3.1:** Overview of the elevation dataset of the Alblasserwaard. A larger map of the elevation dataset can be found in Appendix F.

The other datasets described in this chapter have been selected on their properties and availability. The GBKN is used in chapter 4 for filtered the elevation dataset. The digital palaeogeographic map of the Alblasserwaard and shallow sub-surface drillings are used in the validation process explained in chapter 6.

### 3.1 Principle of Airborne Laser Scanning

Airborne Laser Scanning (ALS) is used to create a Digital Elevation Model (DEM) of an area by sampling individual elevation points. It can be used to efficiently measure large areas with a high point density. The basic system is a laser scanner on board of an airborne platform, for instance a plane or helicopter. The principle of ALS is shown in Figure 3.2. The scanning system sends out pulses in a wide swath underneath the platform.



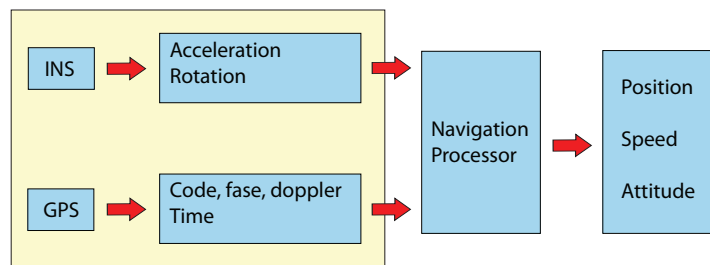
**Figure 3.2:** *The principle of Airborne Laser Scanning.*

The distance (range) between the platform and the ground is measured by emitting a Laser pulse and measuring the travel time  $\Delta T$  (the time for the signal to travel out to the target and back). If we assume that the velocity of light remains constant, than the range  $R$  is derived from:

$$R = \frac{\Delta T}{2} \cdot c \quad (3.1)$$

The location  $(x_i, y_i, z_i)$  of an individual point is calculated using the position, orientation and acceleration of the platform, the direction of the beam and the previously calculated range. To measure these variables different sensors are used; the positioning is done using GPS (Global Positioning System) and the orientation and acceleration by INS (Inertial Navigation System). These basic measurements are combined and processed using a Navigation Processor (NP), often in combination with measurements from one or more fixed GPS reference stations. An overview of this process is shown in Figure 3.3.





**Figure 3.3:** *The different navigation elements used for determining position and attitude.*

The flight altitude, the speed of the plane and the frequency of the laser pulses determine the density of ground points. Flying higher will reduce the density, but flying slower or increasing the frequency of the laser pulses will increase the density. The coordinates of the measurement points are measured in the WGS-84 reference system and converted afterwards to RD (Dutch National Coordinate System) and NAP (Amsterdam Ordnance Datum).

### 3.1.1 FLI-MAP

FLI-MAP (Fast Laser Imaging - Mapping and Profiling) is an Airborne Laser Scanning system mounted to a helicopter, operated by Fugro Aerial Mapping B.V, part of Fugro Geospatial Services. A photograph of the operational system can be seen in Figure 3.4. The booms protruding to the left and the right are the 2 GPS antennas.

The laser scanner in this system is the FLI-MAP 400 VS (Variable Speed) scanner, which is a pulse-scanner combined with a rotating, one-sided mirror. It can be operated at effective pulse rates of 150,000 or 250,000 pulses per second. Effective pulse rate refers to usable pulses, and not the pulse rate of the laser source, which is higher to offset for pulses that are not usable due to scanner mechanics. It is possible that a single pulse is partially reflected by multiple objects, this creates multiple return pulses that can be measured by the scanner. With the FLI-MAP system up to 4 return signals are measured from 1 emitted pulse. The typical absolute precision ( $1\sigma$ ) of a point is 2.5 - 3 cm, but this can vary depending on terrain type and flight altitude.

### 3.1.2 FLI-MAP Data Processing

Different processing algorithms have been applied to the raw point dataset by Fugro. The intension and application of two most important algorithms are described below.

#### Filtering

Because of the high point density and precision it is possible to detect specific features like vegetation and constructions in the raw elevation data. These object can be extracted by using specific



**Figure 3.4:** The FLI-MAP system [18], the laser scanner is indicated with the red ellipse.

filters, various filters have been created by Fugro for this purpose. One of the filters that was applied to the dataset separates ground level and non-ground level points. This leads to the creation of a Digital Surface Model (DSM) of the area.

### Rasterization

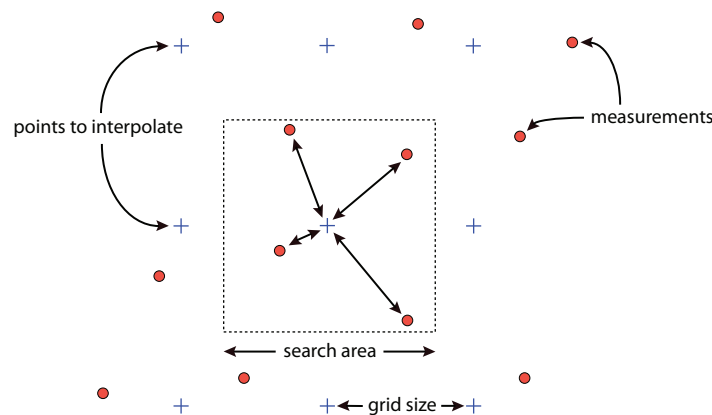
ALS gives an irregular/inhomogeneous point distribution, because measurements are not taken at an equal interval. The raw point cloud can be used for further processing, but due to the sheer amount of points and the size of the dataset this would become problematic. For just the Alblasserwaard the raw dataset accumulates to approximately 2 billion points with a file size of approximately 1 terabyte. For the goal of this research it is not necessary to use all of the data from this raw point cloud. To reduce the amount of points and the processing time, the dataset has been interpolated to a regular, rectangular raster with a grid size of  $0.5 \times 0.5$  m. The interpolation method used is inverse squared distance weighting, shown in Figure 3.5.

The weight ( $h_i$ ) that is given to each of the surrounding points is dependent on the squared distance ( $d_i$ ) to the interpolation point. The interpolated height value ( $\bar{h}$ ) is calculated using equation (3.2).

$$\bar{h} = \frac{\sum_{i=1}^n h_i \times \frac{1}{d_i^2}}{\sum_{i=1}^n \frac{1}{d_i^2}} \quad (3.2)$$

with:

$$d_i = \sqrt{\Delta x_i^2 + \Delta y_i^2}$$



**Figure 3.5:** Schematic overview of the inverse squared distance weighting interpolation method [41]. For AHN-2 the grid size is  $0.5 \times 0.5$  m.

All the processing described in this report will be done based on the raster dataset. In addition to the processing described in the previous paragraphs, the original raster dataset has been converted to tiles of  $2,500 \times 2,000$  pixels ( $1.25 \times 1$  km). In total the Alblasserwaard area is covered by 273 of such tiles. The reason that the data has been divided into tiles is that it makes the files easier to distribute and to manage. Although it does also create challenges in processing.

Processing on this high resolution dataset is very computational intensive. So besides the previously mentioned processing on the dataset the pixels have been resampled to  $2 \times 2$  m. This produces a data product with properties that fits somewhere in between AHN-1 and AHN-2.

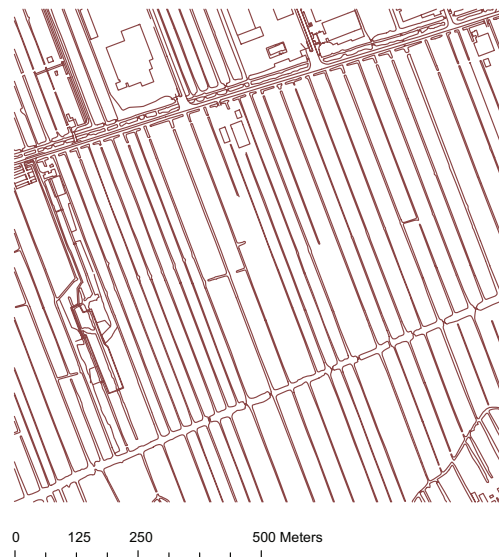
### 3.1.3 Additional FLI-MAP Data

Besides elevation, more information is measured by FLI-MAP. For each individual laser return values are registered for the signal intensity and the colour using RGB values. In addition, FLI-MAP captures aerial photographs and oblique video during the flight.

The additional measurements give information about certain surface properties and it seems unlikely that more information about the subsoil is present in these data products. Although it could be possible that there are differences in the growth of vegetation at the location of buried channel deposits. This would for instance lead to differences in reflectivity. This possibility is not explored in this thesis, but further research could be done to determine if more information can be extracted from the additional FLI-MAP products. However, the use of colour and intensity values is problematic for automated processes like classification/segmentation. This is because there is more variation in these measurements due to for instance different measurement angles.

## 3.2 GBKN

The GBKN (Grootschalige Basiskaart van Nederland) is the Large Scale Standard Map of The Netherlands and is the most detailed and accurate digital topographical database available in the Netherlands [19]. It is unique in the world due to the amount of detail and uniform national coverage. It is a digital map so it is scale-free, but it can be compared to paper maps with a scale between 1:100 and 1:5,000. The accuracy of a point in comparison to another point in the surrounding area is better than 28 cm in suburban areas and better than 56 cm in rural areas [19]. The map is made by a joint venture of different (utility) companies, government organizations and municipalities. It is used mainly for maintenance purposes. The GBKN has a spaghetti-structure: it only contains classified nodes and edges, for instance roadsides, water edges and building contours. An example can be seen in Figure 3.6. Some municipalities have upgraded the map to a surface or object orientated structure. New developments are ongoing to make a fully object orientated topographical database called the TOP10NL [2].



**Figure 3.6:** A random area ( $1 \times 1$  km) of the GBKN in the Alblasserwaard.

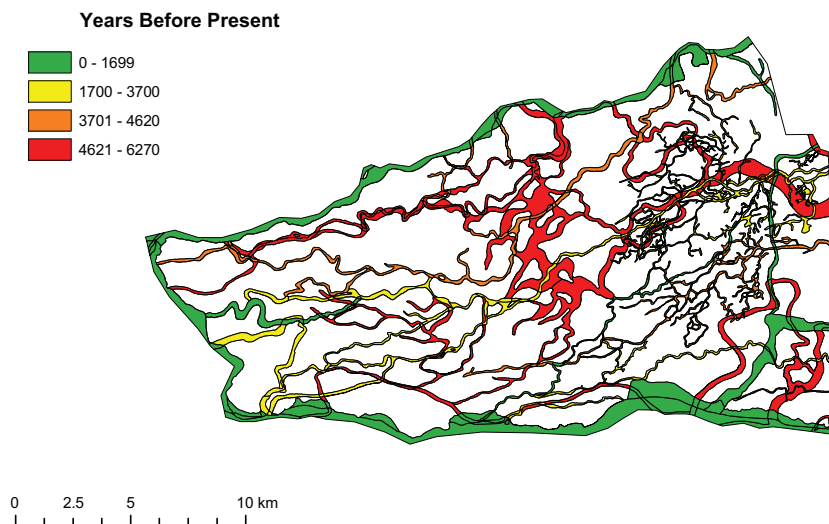
## 3.3 Digital palaeogeographic map

A palaeogeographic map of the whole Rhine-Meuse delta during the Holocene (including the locations of buried channel deposits) was published by Berendsen and Stouthamer [5]. The Alblasserwaard section of this digital map (seen in Figure 3.7) is used as a reference dataset in this research. It is compared to the detection result and the differences between the two maps are analysed to determine why they occur. Questions that will be treated are for example: which age of buried channel deposits can be detected and what are the differences in the location of buried channels deposits between the two maps?

The map is based on more than 25 years of field research using over 200,000 manual boreholes, 45,000 archaeological findings and 1,200 radiocarbon datings. This is a time-consuming process that requires a lot of manual labour, for this reason not the whole Dutch river delta has been mapped in the same amount of detail.

Some properties of the map are:

- It is stored in a vector format, where each individual area consists of a polygon.
- The map contains up to 12 different attributes for each area: size, perimeter length, age, year of beginning, year of ending, etc.
- The age of the channels on the map are given in years Before Present (yBP), where the Present is defined as the year 1950.



**Figure 3.7:** *Palaeogeographic map of the Alblasterwaard containing buried channel deposits [5].*

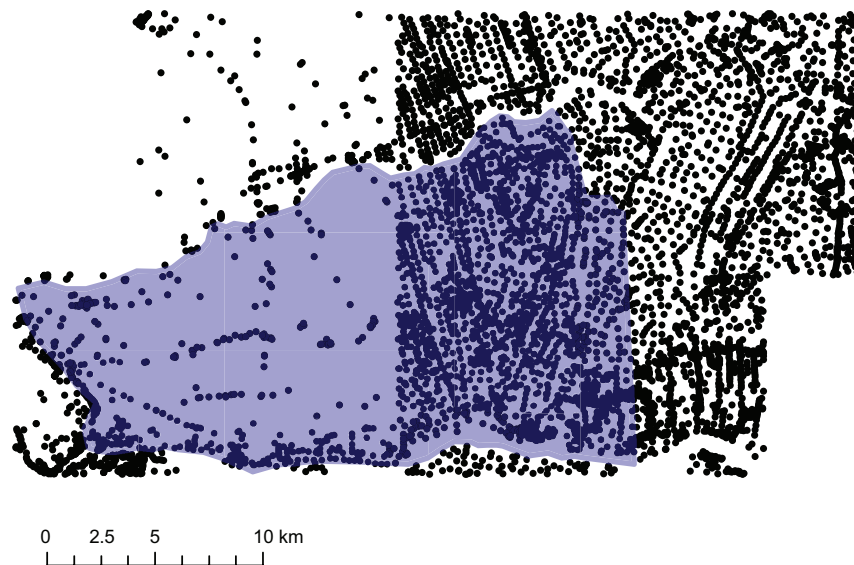
## 3.4 DINOLOKET

The DINO system (Data en Informatie van de Nederlandse Ondergrond) is the central storage site for geoscientific data on the shallow and deep Dutch subsurface. The archive contains shallow and deep drillings, groundwater data, cone penetration tests, geo-electrical measurements, results from geological, geochemical and geomechanical sample analysis, borehole measurements and seismic data.

The shallow drilling measurements from the DINOLOKET are suitable to use as reference data for this research. They cover primarily the shallow subsurface and contain information about the type of sediments and their depth. Classification is performed using official standards, so deposits are described in an unambiguous manner.

The drillings were acquired using two methods: manual and mechanical. With the manual method, it is possible to drill down four to twelve meters, depending on the soil composition. Several manual drillings can be made in a single day to ascertain the position of layers present in the subsurface. Mechanical drillings can reach down hundreds of meters and their execution can take days.

The location of all the drillings in the Alblasserwaard that were available for this research can be seen in Figure 3.8. A division can be made based on different measurement densities: the eastern part has a high density of available drillings and the western part has a very low density. The difference is caused due to data licensing limitations of the DINOLoket, so it was not possible to acquire all the points on the western side. An example of a single drilling can be seen in Appendix C.



**Figure 3.8:** *The black dots indicate the location of the drillings acquired from the DINOLoket. The Alblasserwaard is indicated with the light blue colour.*

# Chapter 4

## Filtering procedure

This chapter describes two filtering methods which are applied to the FLI-MAP data to remove objects like roads, trenches, buildings and water surfaces that do not belong to buried channel deposits. If unaddressed these objects complicate the detection of buried channel deposits as will be described in chapter 5. The filtering method in this chapter uses existing topographical data and terrain characteristics, in this case variability of the terrain.

### 4.1 Filtering using GBKN

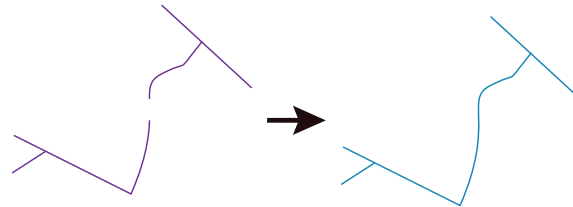
The GBKN has been introduced in section 3.2. The GBKN has a line-oriented data structure, this type of data can not directly be used for filtering purposes. A surface or object orientated map of the Alblasserwaard is needed so that an overlay procedure can be applied to remove objects in the GBKN from the FLI-Map data. To achieve this some processing has to be done on the GBKN data to convert this line dataset into a format consisting of area-objects. To create an area-object the individual lines have to be automatically connected, merged and converted into surfaces. However, there are errors in the database: lines sometimes do not connect exactly or lines intersect without a node. Such situations have to be removed or adapted so that they will not negatively influence the process of area-building. As a final step in the procedure the area closely surrounding the objects is added to the surface that is to be removed, due to the fact that it is highly likely that changes have been made. All the steps mentioned in this paragraph will be explained in more detail in the next paragraphs.

#### 4.1.1 Snapping

Many lines in the GBKN do not exactly connect to each other, often due to measurement errors, inaccuracies, time difference between measurements or the combination of different measurement techniques. This poses a problem, because area objects can only be made from non-connecting lines. To solve this problem the end of an unconnected edge has to be attached to the nearest line or point



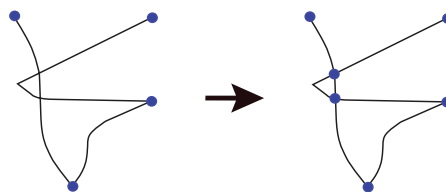
that lies within a certain distance/tolerance, this operation is called snapping and can be seen in Figure 4.1. In this research a maximum distance of 10 cm has been used for snapping.



**Figure 4.1:** *The result of the Snapping operator (adapted from [33]).*

### 4.1.2 Self Intersection

If the lines are eventually going to be converted to polygons, they are not allowed to intersect or self-intersect without a node at the intersection point (Figure 4.2). This has to be checked and rectified for every feature. If a line self-intersects, it is split into separate lines with the same attributes and a node is added at the intersection point.



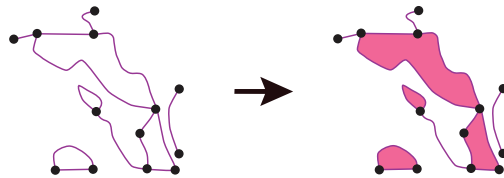
**Figure 4.2:** *Example of an intersection and the corresponding solution (adapted from [33]).*

### 4.1.3 Area building

After cleaning up the database using the snapping and intersection operations the next step is to form areas from the lines. This is done by combining connected lines to each other, eventually forming area features, as can be seen in Figure 4.3.

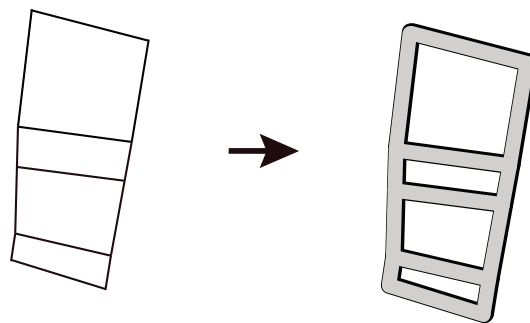
### 4.1.4 Buffering

The surface within a certain distance of an object will in nearly all cases also have been disturbed. For instance material that has been dredged out of a channel will have been deposited alongside it, roads



**Figure 4.3:** *The result of the area building operator (adapted from [33]).*

have been elevated or leveled out completely. It is hard to put an exact number on the distance that should be buffered around objects, but an estimation based on the FLI-MAP dataset is that a distance of 3 meters is adequate in most situations. This distance is still quite conservative, but it must be taken into account that not too much of the dataset is filtered out for it may also contain information that can be used to detect buried channel deposits. An added effect of the buffer-operation is that lines that have not been formed into an area-object in the previous processing steps can still be used to filter out unwanted objects.



**Figure 4.4:** *The result of the buffering operator (adapted from [33]).*

#### 4.1.5 Further processing

The result of the previously explained operations is a shape-file where the GBKN line features have been transformed into areas including the local surrounding area (Figure 4.5). However different errors in the database and effects of the processing resulted in complete fields which were incorrectly included. If ignored this would lead to these fields being removed from the dataset. This would be unfavourable due to the fact that most of the information about buried channel deposits can be extracted from the elevation of the fields. Two methods were applied to make sure that these fields were removed from the GBKN area dataset and thus remained in the elevation dataset.

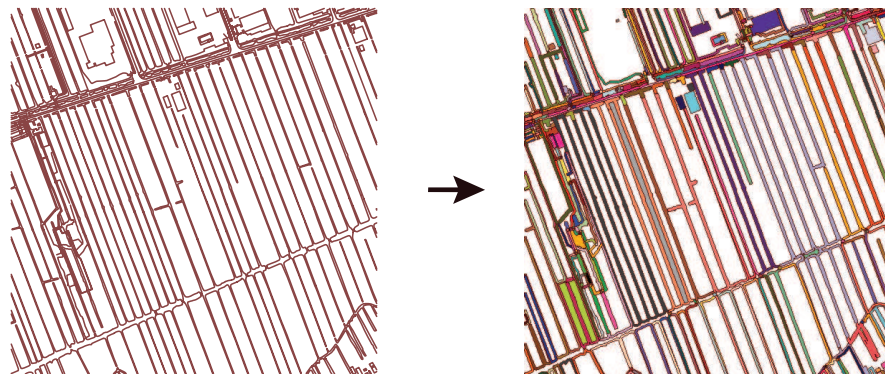
The first method is a selection operation based on area size and spread. The fields that need to be removed are relatively large compared to all the other objects and mostly square or rectangular, so they are compact. For a measure of spread the surface area of the object is divided by the surface area of the smallest enclosing rectangular bounding box. This leads to an indication of how compact or spread out the object is.

Decisions were based on the following rules:

Object < 4,000 m <sup>2</sup>	→	→	→	<b>keep</b>
Object > 4,000 m <sup>2</sup>	→	spread < 0.15	→	<b>keep</b>
Object > 4,000 m <sup>2</sup>	→	spread > 0.15	→	<b>remove</b>

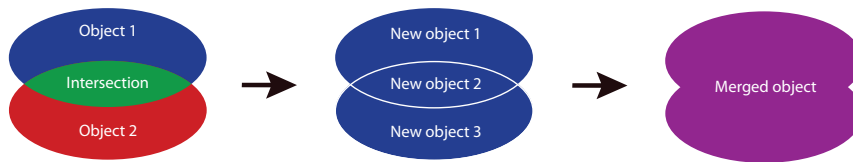
The objects with an area smaller than 4,000 m<sup>2</sup> remain in the database. For objects with an area larger than 4,000 m<sup>2</sup> the spread of each object has been calculated. If the spread is lower than 0.15 than this object also remains in the database, if the spread is higher than 0.15 than the object is removed. So objects which are very compact are removed, but objects which are very spread out are maintained. The values of 4,000m<sup>2</sup> and 0.15 have been chosen by applying trial and error, it turned out in practice that these numbers gave a good separation.

The thresholds have been chosen somewhat conservatively, so no objects were removed that should not have been removed. This also implies that fields that should have been removed were not. To solve this, a visual inspection was performed to remove the remaining fields that had not been removed automatically. In total about 150 fields had to be manually removed from the area.



**Figure 4.5:** *The initial GBKN dataset (left) and the final result (right) after processing. Line-objects are converted into area-objects, so it can be used for filtering. In this figure each area object has its own colour.*

For the Alblasserwaard 697,283 area-objects were created using the previously explained procedures; this large amount of object makes it impossible to use it efficiently in the filtering procedure. To solve this problem, all small area-objects that (partially) overlapped need to be intersected to form objects that do not overlap. These object were then merged together to form one very big shape file. The process is shown in Figure 4.6. The resulting single shape-file could then be efficiently used to filter out topographic objects that are present in the GBKN.

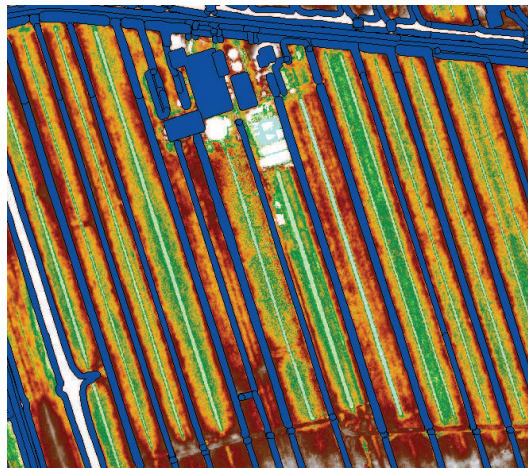


**Figure 4.6:** An example of the intersection and merging of overlapping objects.

Despite this filtering method unwanted features still remain present in the FLI-MAP data; for instance small trenches and uncharted objects. These features are not registered in (topographical) databases so a different approach has to be taken if they are to be filtered out. Further steps in the removal of these objects are described in the next paragraph.

## 4.2 Variability filter

Not all objects and terrain features are registered in the GBKN; some objects are too small or not important enough, while in other cases no differentiation has been made. This first case can be observed in Figure 4.7 where small trenches have been dug to prevent accumulation of precipitation, these trenches are not registered so they can not be removed by overlaying a dataset. Removal of these remaining features must be done by using the existing terrain properties to determine if they can be part of buried channel deposits.



**Figure 4.7:** In this figure the dark blue object is the GBKN mask, the white areas contain no data. Small features can be observed that are not covered by the GBKN mask. For instance the light blue channels do not contain water but are for drainage purposes, these channels are not registered in the GBKN.

### 4.2.1 Variability

The natural tendency of a peat/clay/sand surface is that it is smooth so it has a relatively low variability; if this was not the case it means that something of someone has disturbed it. These disturbances are for instance trenches, buildings and other artificial objects; they cause a high local variability. Based on this assumption the variability can be used to filter out unwanted objects from the height data, leaving smooth surfaces like grasslands and agricultural fields where buried channel deposits can be detected. The variance ( $\sigma^2$ ) is calculated using a window that moves over the height model. This is done by applying equation (4.1). All the height values ( $h_i$ ) in the windows are used and  $n$  is the number of height values.

$$\sigma^2 = \frac{1}{n-1} \sum_{i=1}^n (h_i - \bar{h})^2 \quad (4.1)$$

where the average height ( $\bar{h}$ ) of the window is calculated by using:

$$\bar{h} = \frac{1}{n} \sum_{i=1}^n h_i \quad (4.2)$$

A small window can be used to detect local effects, large scale effects can be detected by using a large window. The disturbances that we would like to filter out are very local, so it was opted to use a relatively small window of  $3 \times 3$  pixels. A window this size still has enough pixels to calculate a variance. An example of a moving variance filter on an example dataset is given below:

Example dataset:

2	2	3	3	1	4	7
2	2	2	4	4	6	1
2	2	3	1	5	4	9

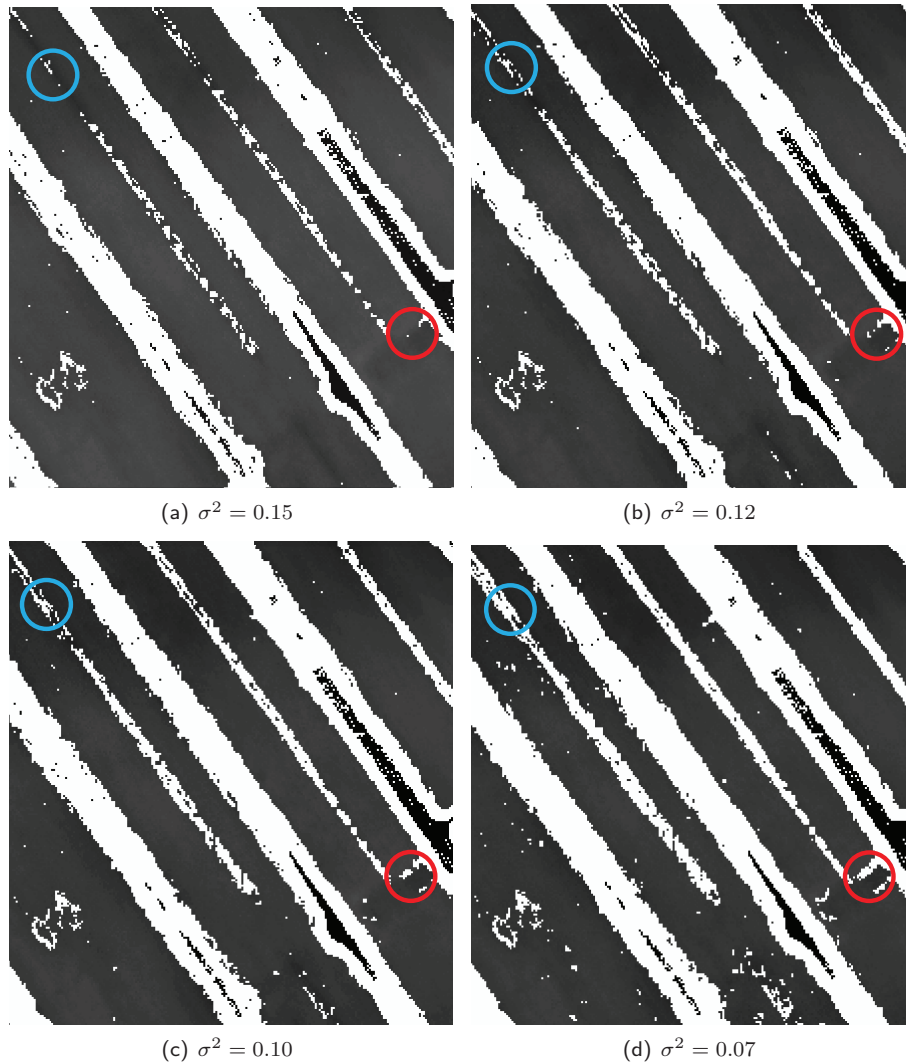
Moving variance window:

$h_1$	$h_2$	$h_3$
$h_4$	$h_5$	$h_6$
$h_7$	$h_8$	$h_9$

Output of the moving variance window, using equation 4.1:

0.19	0.78	1.86	2.78	6.00
------	------	------	------	------

In this filtering step the assumption was made that the variance in a certain neighbourhood should not become too high. The areas which do not belong to buried rivers are in that case selected by applying a threshold at a certain variance. In the Figures 4.8 different thresholds have been applied.



**Figure 4.8:** The result of 4 different variance thresholds applied to the same area (in RD coordinates:  $x = 116,370$   $y = 436,876$ ). Inside of the blue and red circle the differences between the results of using different thresholds can be observed.

Applying a variance threshold of  $\sigma^2 = 0.15$  leaves most of the data intact. Each time the threshold is lowered more data are filtered out. At a variance threshold of  $\sigma^2 = 0.12$  and  $\sigma^2 = 0.10$  it can be seen that some of the small trenches are removed, because they cause a higher local variance compared to the surrounding flat area. At a variance threshold of  $\sigma^2 = 0.10$  also points from the fields start to be filtered out, this indicates that this threshold value is too low. For the actual filtering process a variance threshold of  $\sigma^2 = 0.10$  was applied, because this threshold value removes unwanted points, but leaves most of the points that could contain information about the location of buried channel

deposits.

### 4.3 Density filter

After applying the previous two filters, a large part of the unwanted objects were filtered out. However, some distinguishable objects remained in the dataset, these are for instance buildings, front- and backyards, driveways and centre dividers (the area between opposite road lanes). A possibility explored here is to remove these areas by applying a filter based on local point density.

The density threshold is the minimum density which should be present inside the moving window, for example 40% means that inside the moving windows at least 40% of the pixels should have a height value. This means that at most 60% of the pixels are empty (transparent in all the figures).

There are 2 different variables that can be varied when applying a density filter: the windows size and the density threshold. The effects of varying both variables when applying a density filter on the dataset have been studied. The effects are given in table 4.1.

**Table 4.1:** *Effects of windows size variation versus different density thresholds using a density filter.*

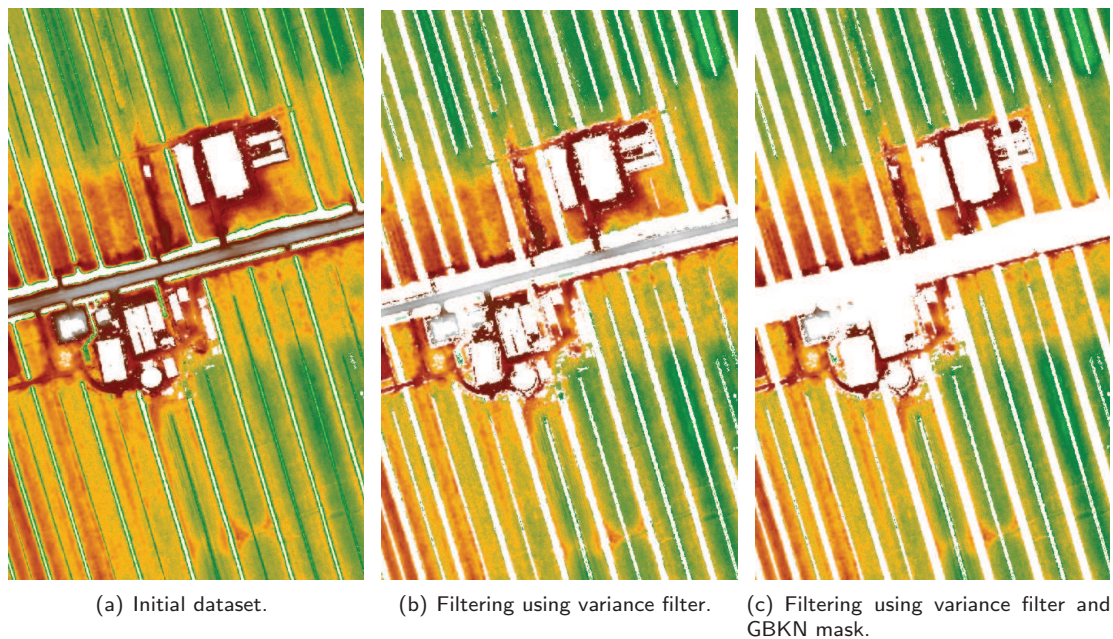
		Window Size		
		31x31	41x41	51x51
Density	40%	borders of fields removed	not enough removed	not enough removed
	45%	borders of fields removed	not enough removed	not enough removed
	50%	borders of fields removed	few borders of fields removed	few borders of fields removed
	60%	insides of fields removed	insides of fields removed	borders of fields removed

The conclusion is that both variables have different influences on the final result of the density filter. There is no ideal solution in this case, but the best possible relation between the variables is a windows size of 41x41 pixels with a density threshold of 50%.

### 4.4 Results

An example of the results obtained during the filtering procedure can be seen in Figure 4.9. Most of the large artificial objects have been removed; roads, the larger trenches, some of the smaller trenches and parts of the building. This area is representative of what happens throughout the entire Alblasserwaard when the filtering procedure is applied.





**Figure 4.9:** Results obtained during the filtering procedure.

#### 4.4.1 Remarks

If a close look is taken at the dataset certain effect can be observed that could lead to errors, they can be subdivided into different types or effects:

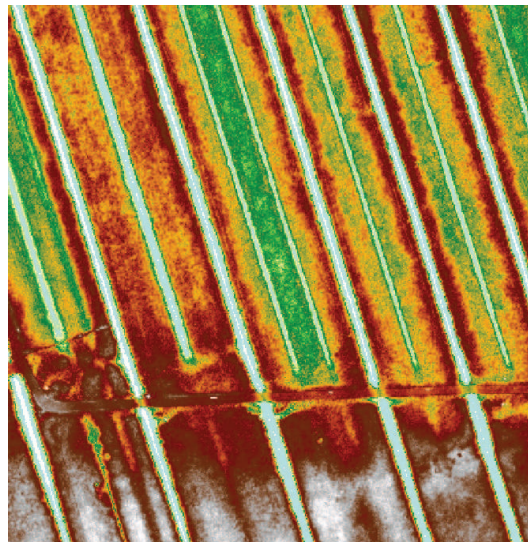
##### **Disturbance of the signal of buried channel deposits:**

- **Plowing**  
Heavy machinery used by farmers to plow fields cause material to be redistributed over a larger area, in fact causing a smoothing of the area. Also the principal direction of the plowing can be discerned in many cases by looking at the direction of the smoothing.
- **Leveling/equalization**  
In a few cases individual farmers have leveled whole fields, adding material to raise the surface level. Although in some cases material was removed (possibly the underlying sand was dug away). Both situation lead to a near total loss of the signal of buried channel deposits.
- **Ground water level differences**  
Differences in ground water level between polders may cause abrupt elevation changes at the boundary of these polders. Due to this buried channel deposits seem to suddenly stop when reaching these boundaries.



**Introduction of signals that resemble buried channel deposits:**

- **Deposition of dredged material**  
During maintenance of trenches, dredged material is often deposited besides the trench onto the fields. This effect causes an increase in the heights of the area around the trench (seen in Figure 4.10). This is the reason why no interpolation algorithm is applied to fill in the gaps caused by the filtering process: this would lead to trenches that are relatively elevated in comparison to the surrounding area.
- **Filled in trenches**  
Due to agricultural consolidation old trenches became redundant and were filled in with material. Due to overfilling, underfilling or compaction, elevation differences remain. Remnants of such processes can still be seen in certain areas in the FLI-MAP dataset, in Figure 4.10 parts of the original trench can for instance still be seen. This effect is compensated for a large part by the GBKN. Due to the fact that the GBKN has not been updated in this area for quite some time, the old situation with the old trenches is still represented in the dataset. This results in the removal of most of the old trenches, although sometimes there have been some updates to the GBKN and then the trenches are not removed.
- **Vegetation**  
In some areas it appears as if whole fields have been raised. This is caused by the Lidar data reflecting on dense/high crops, like for instance corn or wheat.



**Figure 4.10:** *Example of a filled in trench and deposition of dredged material near the sides of trenches.*

#### 4.4.2 Discarded filters

A number of different filters was tested but not applied in the final procedure. This is due to different reasons, which are explained below for each filter.

**Absolute height filter:**

The surface level varies throughout the Alblasserwaard, this can also be seen in the different waterlevels that are maintained in the Alblasserwaard (between -0,80 m and -2 m). Application of this filter would lead to the correct filtering in one area, but to the incorrect removal of data in other areas.

**Relative height filter:**

Remnants of the filtering step are mostly in the same order of height (decimeters) as buried channel deposits. It is not possible to make a distinction between these two classes only based on height, so application of this filter would remove unwanted objects as well as height data containing information about buried channel deposits.

**Local gradient filter (windows size of 2x2 pixels or 3x3 pixels):**

This filter gave approximately the same result compared to the variance filter. No additional areas were removed, so there is no need to apply this filter.

**Fourier filter:**

The data were converted into the frequency domain using a 2D Fourier transformation. In the frequency domain the spectrum of the higher frequencies was removed, the middle and low frequencies were preserved. An inverse 2D Fourier transformation was performed to transform the data back into the spatial domain (the image). Looking at the result, this method did not only filter but also introduced a lot of new errors.



# Chapter 5

## Detection

The objective of this thesis is the automatic detection of buried channel deposits from high resolution laser altimetry data (FLI-MAP). In this chapter a method is described for the detection of these channels deposits using techniques from image-processing and remote sensing. An overview of the workflow of the entire detection process is given in Appendix A.

The data used for the detection of buried channel deposits in this research is geo-referenced height data from the FLI-MAP system. Additional information like Returned Signal Intensity is measured with FLI-MAP, but this information was not available for this research as explained in chapter 3.1.3. To use as much of the available information contained in the height data in the detection process, four derived attributes are calculated in section 5.1 by using the elevation of surrounding points. It is explained how these attributes are calculated, what their characteristics are and in what way these characteristics contribute to the detection process. By combining these attributes we obtain a multi-band image, which allows the application of multispectral image classification techniques. The structural attributes and the result of the classification will be shown in chapter 6.

In section 5.2 a classification is performed to determine for each point if it belongs to a predefined class. Analysis is performed to determine if 2 or 3 classes should be used. The advantages and disadvantages of this method will also be explained.

Finally, two additional methods for automating the detection process are described. In section 5.3 the implementation and effects of the downsampling of data are analysed. Section 5.4 explains a method to resolve edge effects on tiled datasets.

### 5.1 Structural Attributes

If we look at a single point in the FLI-MAP dataset, this point only contains a height value. However, the height values surrounding this point also hold information that can be related to that specific location. For example the slope of the terrain can be calculated for a single point by using the surrounding height values and their location. An attribute obtained by performing an operation that

uses the surrounding points is called a structural attribute. In this research a total of four structural attributes are calculated (see Figure 5.1), these are: slope, curvature, Topographic Position Index (TPI) and smoothed TPI. For each of these attributes the objective and calculation method will be explained in the next paragraphs. The information from these attributes is used by the classification algorithm for the detection of buried channel deposits.

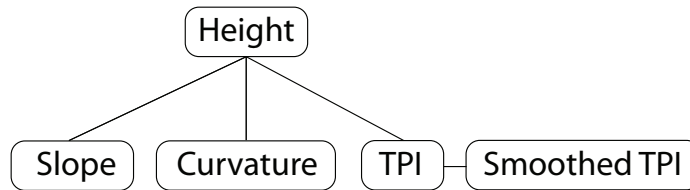


Figure 5.1: Structural attributes derived from the elevation data.

### 5.1.1 Attribute 1 - Slope

Figure 5.2 shows a schematic cross-section of the surface elevation at the location of a buried channel (shown in black). In this illustration small patches of the surface can locally be represented by a first-degree trend surface, also known as a plane (shown in blue). Due to the local topography some of these planes are not horizontal, but lie at a different angle compared to a horizontal surface. The angles between the plane and the zenith are noted with  $\beta$ . If buried channel deposits occur, the surface topography shows a 'bulge' and on both sides this will result in a plane with a gradient. On the top of the 'bulge' the surface is horizontal.

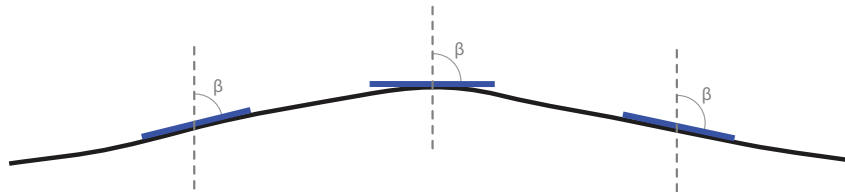


Figure 5.2: A schematic cross-section of a buried channel in the FLI-MAP dataset, including three fitted planes.

This example can be converted to a 3-Dimensional model, where the plane is a surface area. The data values  $(x_i, y_i, z_i)$  can directly be taken from the FLI-MAP dataset. The assumption is that the  $z$ -component of the data is dependent on the  $x$ - and  $y$ -components. So if we want to estimate the height values ( $z$ ) based on the coordinates ( $x$ ) and ( $y$ ), the equation of a plane can be written like:

$$z = c_1 \cdot x + c_2 \cdot y + c_3 \quad (5.1)$$

The parameters  $(c_1)$  and  $(c_2)$  represent the slope of the plane in the perpendicular  $x$  and  $y$  directions.

Because of this, these parameters also contain directional information about the plane. For example if  $c_1 = 0$  and  $c_2 = -0.3$ , the plane is tilted in the  $y$ -direction and not in the  $x$ -direction, if  $c_1 = 0.2$  and  $c_2 = 0.2$  the plane is tilted in both directions. This concept is called 'aspect' and is the direction that the plane faces in relation to the coordinate system of the dataset.

The shape of the patch used for plane fitting is symmetric. Symmetric neighbourhoods are used when no changes in signal are expected with changes of direction. This is also known as isotropy. The location and direction of buried channel deposits are not known beforehand and the use of a directional neighbourhood would make it more difficult to detect channels that are lying in a different direction. So for calculating this attribute a symmetric neighbourhood is used.

The fitting of the plane was done using Least-Squares estimation [37]. This determines the plane parameters  $c_1$ ,  $c_2$  and  $c_3$  such that the sum of the squared errors between the height values and the plane is minimized. Note that the errors are calculated in the  $z$ -direction and not perpendicular to the plane.

The model of observation equations used for determining the plane parameters is:

$$E\{z\} = \mathbf{A}x \quad (5.2)$$

Here  $z$  is the vector of observations,  $\mathbf{A}$  is the model matrix and  $x$  is the vector of model parameters. In this research the number of observations was in the order of:  $2,000 \leq n \leq 2,500$ .

$$\begin{bmatrix} z_1 \\ z_2 \\ \vdots \\ z_n \end{bmatrix} = \begin{bmatrix} x_1 & y_1 & 1 \\ x_2 & y_2 & 1 \\ \vdots & \vdots & \vdots \\ x_n & y_n & 1 \end{bmatrix} \begin{bmatrix} c_1 \\ c_2 \\ c_3 \end{bmatrix} \quad (5.3)$$

Assuming that the weights of all the points are equal, this problem can be solved using:

$$\hat{x} = (\mathbf{A}^T \mathbf{A})^{-1} \mathbf{A}^T z \quad (5.4)$$

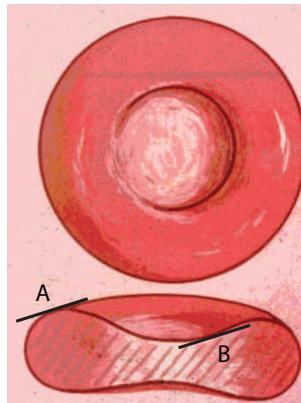
where  $\hat{x}$  is the vector containing the parameters  $(\hat{c}_1, \hat{c}_2, \hat{c}_3)$  of the fitted plane. The symbol  $(\mathbf{A}^T)$  denotes the transpose of matrix  $\mathbf{A}$ . The slope of the plane is calculated by using the parameters  $\hat{c}_1$  and  $\hat{c}_2$ . Using Horn's method [10], the slope ( $s$ ) is given by:

$$\begin{aligned} \tan s &= \sqrt{(\partial z / \partial x)^2 + (\partial z / \partial y)^2} \\ s &= \tan^{-1} \sqrt{\hat{c}_1^2 + \hat{c}_2^2} \end{aligned} \quad (5.5)$$

The slope of the fitted plane is in degrees, ranging between  $0^\circ$  and  $90^\circ$ . Using Horn's method, the directional information (aspect) of the plane is lost. This is not a problem because this detection algorithm does not use the aspect information. Research could be done to determine if this information could be used in a different type of detection algorithm.

### 5.1.2 Attribute 2 - Curvature

The first order derivatives of elevation are slope and aspect; these have been explained in the previous paragraph. They give information about the terrain but do not directly specify its shape. To obtain more information about the shape of a terrain you can calculate the curvature, which is a measure of how curved a surface is within a specified area. Convex is another word for curving out or bulging outwards, the opposite of convex is concave, which means curving inwards. In a convexity map a positive curvature indicates that the surface is convex at that area. A negative curvature indicates that the surface is concave, and a value of zero indicates that the surface is flat. Note that a flat surface with zero curvature can have a slope, so it does not have to be a horizontal surface. If we look at a human red blood cell for example (Figure 5.3), at the centre it is concave (has a negative curvature), at the same time the sides are convex (positive curvature). If we look in more detail there are certain points where the slope is the same, for instance point A and B, but there is a difference in curvature. Seen from the outside, point A lies on a convex surface and point B on a concave surface.



**Figure 5.3:** A schematic drawing of human red blood cells, seen from the top and the side. Point A and B have the same slope, but different curvature values.

If we now look at the way buried channel deposits are visible at the surface, in general we can say that they are convex because they 'bulge' outwards. This means that the structural attribute curvature can be used for detection purposes.

Two common ways to calculate curvature are 'mean curvature' and 'Gaussian curvature'. In this research the mean curvature is used instead of the Gaussian curvature. This is because buried channel deposits are mostly curved in one direction and the Gaussian curvature is more suitable for detecting curvatures in two directions. The mean curvature ( $H$ ) is the mean of the two principal curvatures:  $k_1$  (maximal curvature) and  $k_2$  (minimal curvature). At a given point of a surface these two principal curvatures measure how the surface bends by different amounts in different directions at that point. The principle curvature directions are perpendicular to each other.

From [34] the equation of the mean curvature can be written as:

$$H = (k_1 + k_2)/2 \quad (5.6)$$

It is common to calculate the mean curvature and the Gaussian curvature first and get  $k_1$  and  $k_2$  by inversion. So in this case we only need to apply the direct way of calculating  $H$ . This is done using equation 5.7 from [7]. Here  $z_x$ ,  $z_y$ ,  $z_{xx}$ ,  $z_{yy}$  and  $z_{xy}$  are the partial derivatives of the height  $z$  with respect to the directions  $x$  and  $y$ .

$$H = \frac{z_{xx}(1 + z_y^2) + z_{yy}(1 + z_x^2) - 2z_x z_y z_{xy}}{2(1 + z_x^2 + z_y^2)^{3/2}} \quad (5.7)$$

If we want to calculate  $H$  we need to estimate the partial derivatives based on the height data. There are various options to do this, but the method used for this research is to fit a 2-Dimensional polynomial through a selected patch of height data using Least Squares estimation. A second degree surface whose 6 functions are orthogonal is given in equation 5.8. If it is assumed that the neighbourhood of a patch is squared and it contains  $L \times L$  measurements, then the horizontal and vertical distance from the centre pixel to the side of the patch is:  $M = (L - 1)/2$ . With  $\Delta x$  the  $x$ -distance between the centre pixel and the calculation point and  $\Delta y$  the  $y$ -distance between the centre pixel and the calculation point.

$$z = d_1 + d_2 \Delta x + d_3 \Delta y + d_4 ((\Delta x)^2 - M(M + 1)/3) + d_5 \Delta x \Delta y + d_6 ((\Delta y)^2 - M(M + 1)/3) \quad (5.8)$$

Again the model of observation equations is used (see eq. 5.2), and implemented [28]. This results in equation 5.9. The height of the points are in the vector of observations ( $z$ ), the model parameters in the model matrix ( $\mathbf{A}$ ) and the parameters for the partial derivatives are in the vector of model parameters ( $x$ ).

$$E\{z\} = \mathbf{A}x$$

$$\begin{bmatrix} z_1 \\ z_2 \\ \vdots \\ z_n \end{bmatrix} = \begin{bmatrix} 1 & \Delta x_1 & \Delta y_1 & ((\Delta x_1)^2 - M(M + 1)/3) & \Delta x_1 \Delta y_1 & ((\Delta y_1)^2 - M(M + 1)/3) \\ 1 & \Delta x_2 & \Delta y_2 & ((\Delta x_2)^2 - M(M + 1)/3) & \Delta x_2 \Delta y_2 & ((\Delta y_2)^2 - M(M + 1)/3) \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 1 & \Delta x_n & \Delta y_n & ((\Delta x_n)^2 - M(M + 1)/3) & \Delta x_n \Delta y_n & ((\Delta y_n)^2 - M(M + 1)/3) \end{bmatrix} \begin{bmatrix} d_1 \\ d_2 \\ d_3 \\ d_4 \\ d_5 \\ d_6 \end{bmatrix} \quad (5.9)$$

This problem is solved by applying equation 5.4 again. From the parameters  $\hat{x}$ , the partial derivatives for the point at the centre of the window are then derived as:

$$z_x = \hat{d}_2, \quad z_y = \hat{d}_3, \quad z_{xx} = 2\hat{d}_4, \quad z_{xy} = \hat{d}_5, \quad z_{yy} = 2\hat{d}_6. \quad (5.10)$$

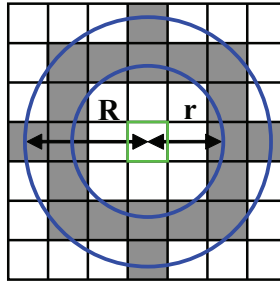
These partial derivatives are then inserted into equation 5.7 to calculate a mean curvature value ( $H$ ) for the centre pixel of the patch. This patch moves across the dataset so a mean curvature value can be calculated for every location in the dataset.



### 5.1.3 Attribute 3 - TPI (Topographic Position Index)

The main characteristic of buried channels is that they have a higher elevation compared to the surrounding surface area. These height differences are in the order ranging from centimeters up to meters. To make use of this characteristic an algorithm is needed that takes the elevation of a single point and compares it to the elevation of the area or neighbourhood around it. A measure for this is the Topographic Position Index (TPI). TPI is a measure of the elevation of a location compared to the surrounding landscape [44]. If you take an arbitrary point on a surface and compare it to the surrounding topography, it may be part of a plain, a hilltop, a valley or another feature. Using TPI these various landscape classes can be distinguished when automatically classifying a region. TPI has already been used for detecting and classifying buried channels from unfiltered AHN-1 raster-data [40]. However, classification was only based on the TPI value and the local slope.

To obtain the TPI-value of a single pixel the difference between its elevation and the average elevation of the neighbourhood around that cell is calculated. The most frequently used shape for the TPI neighbourhood is annular, so all cells within a certain range are used in the calculation. This is a major difference to the plane and curvature fitting where a square neighbourhood is used. An example is shown in Figure 5.4.



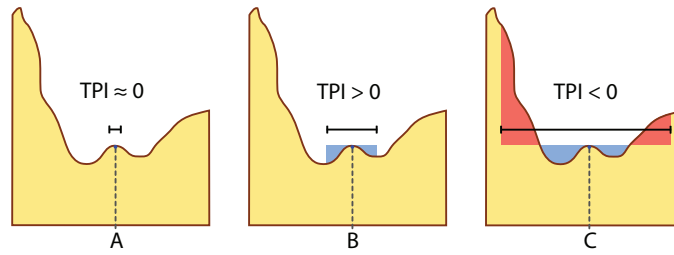
**Figure 5.4:** Example of a TPI neighbourhood to calculate the TPI value of the green centre pixel using the average height value of the gray pixels (Modified from [44]). The annulus is defined by the radius of the inner circle ( $r$ ) and the radius of the outer circle by ( $R$ ).

Using an annular neighbourhood, the TPI-value ( $\tau_{tpi}$ ) is calculated by subtracting the height of the centre pixel ( $z_{tpi}$ ) by the average of the height values that are between range  $r$  and  $R$ :

$$\tau_{tpi} = z_{tpi} - \frac{\sum_{i \in N} z_i}{|N|} \quad (5.11)$$

$$N = \{(x_i, y_i, z_i) : r < \|(x_i, y_i) - (\bar{x}, \bar{y})\| < R\} \quad (5.12)$$

A positive TPI-value means that the cell is higher than its surroundings (at the specified neighbourhood size) while negative values mean it is lower. A TPI-value of zero or close to zero means that either the cell lies on a flat area or that it lies on a constant slope.



**Figure 5.5:** TPI values at 3 different scales, indicating the scale dependency of TPI (modified from [44]).

TPI is very scale-dependent, as can be seen in Figure 5.5. The same point may look like a flat plain to an insect (A), look like a rocky outcrop to a human walking past it (B) and look like the bottom of a valley to a bird flying above (C). The values calculated by this algorithm depend entirely on the scale that is used to analyze the landscape; by adjusting the size of the neighbourhood different features can be extracted. Determining the optimum size of the annulus is treated in paragraph 5.3.1.

#### 5.1.4 Attribute 4 - Smoothed TPI

The result of the TPI operation is still very detailed because it calculated the height difference between single height points and the average surroundings. In this way small objects continue to influence the results, causing misclassifications of these areas.

To make a distinction between smaller topographic features and a larger topographic feature like the buried channel deposits, an averaging operation is performed on the result of the TPI operation. A window is used to select points around a centre pixel, all at most a distance ( $v$ ) away in horizontal or vertical direction. The average TPI value for this point ( $\overline{\tau_{tpi}}$ ) is the average of the heights of the selected points ( $N$ ) that fall inside the window.

$$\overline{\tau_{tpi}} = \frac{\sum_{i \in N} \tau_{tpi}}{|N|} \quad (5.13)$$

$$N = \{(x_i, y_i, \tau_{tpi}) : (x_{i-v} < x_i < x_{i+v}) \wedge (y_{i-v} < y_i < y_{i+v})\}$$

Care must be taken that the averaging area is not too large, otherwise smaller buried channel deposits may possibly be removed. The size of the averaging window is treated in paragraph 5.3.1.

## 5.2 Classification

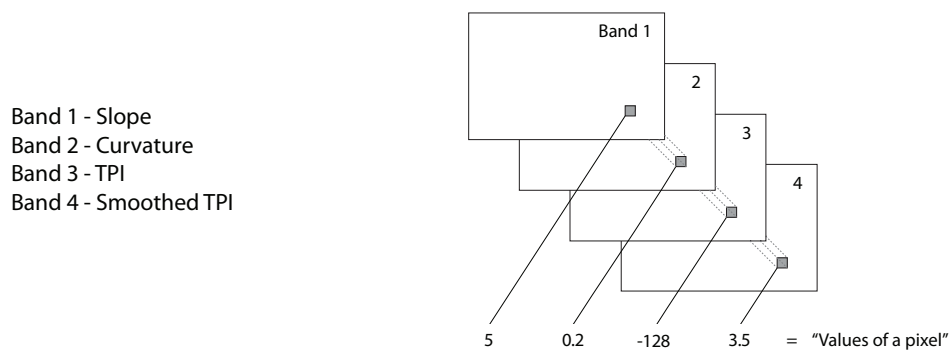
The process of classification is used in Remote Sensing to determine to which predefined class a pixel or an area belongs. For this research we can not use a straightforward classification, such as

classification based on fixed thresholds for elevation, slope or TPI. If a classification is based on the absolute height of an area, a number of problems occur: deeper channels deposits that cause less surface height elevation would in some cases not be classified as buried channels. Some elevated artifacts like roads still remain in the dataset, this despite the filtering process explained in chapter 4. The most important reason why a uniform elevation cut-off can not be used is that the average height for different areas (and for each tile) is very different throughout the Alblasserwaard.

Various combinations of criteria and thresholds were tried for classification, but none led to a satisfying result. Therefore it was decided to apply more robust and advanced techniques for the detection of buried channel deposits. In this paragraph a suitable classification method is explained and applied.

### 5.2.1 Multispectral image classification

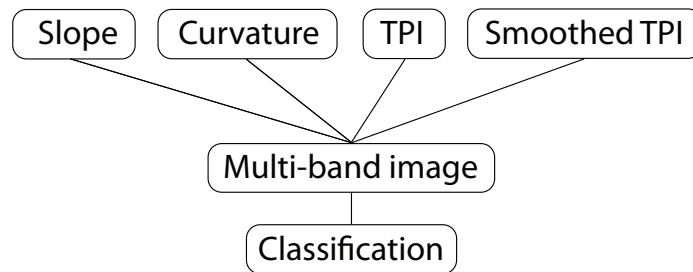
A definition of multispectral image classification is: "The process of assigning individual pixels of a multispectral image to categories on the basis of spectral reflectance characteristics" [12]. Multispectral image classification originated from remote sensing. It is used to classify areas where different reflecting wavelengths are measured at the same time. This results in the creation of multiple images of the same area. In this research no spectral reflectance bands were used, however the previously calculated structural attributes can be used as bands. The original height data are not used as a separate band. This is because of the fluctuation of the average height throughout the Alblasserwaard. This problem does not occur with the structural attributes because they are all relative (based on the surrounding points). The bands used in this classification are graphically represented in Figure 5.6:



**Figure 5.6:** An example of a single pixel and each of its values in the separate bands (picture adapted from [12]).

An advantage of this approach is that information contained in these separate bands can all be used at the same time for classification. Furthermore, no fixed criteria or threshold values are required due to the fact that more advanced classification methods are available. And finally, image classification is a well documented process which is repeatable, verifiable and removes much of the human interpretation done by the operator. Another advantage is that bands can be added or removed, so if new structural attributes are selected or developed they can easily be incorporated. A possibility that is not further explored in this thesis is that the same structural attribute could be calculated a number of times using different window sizes. This is done to obtain information about features that have different operation scales, for instance a buried channel with a smaller side channel.

The structural attributes are combined to form a multi-band image. The previously mentioned image classification methods can be applied on this image. An overview of the classification method up to this point is given in Figure 5.7.



**Figure 5.7:** *Structural attributes used in the classification process.*

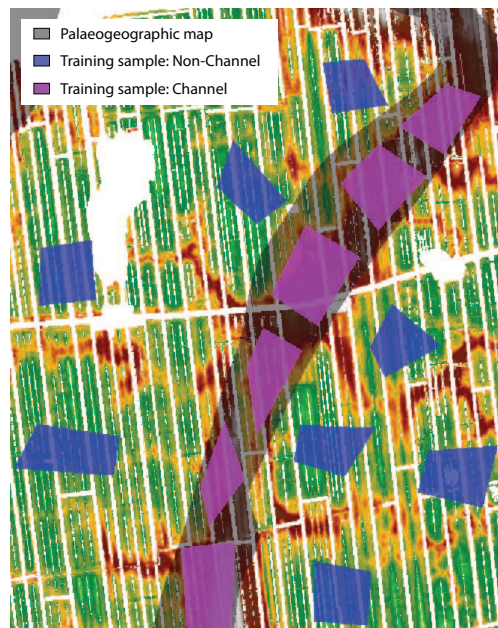
The process of classification is performed by assigning each pixel to a class based on the values from the different bands. Different methods of assigning a pixel to a class are available, each with their own properties. For this research the classification method used was a Maximum Likelihood Classifier [36]. This is a well known and documented classification method which is often used on Remote Sensing data.

The Maximum Likelihood Classifier uses statistics from class signatures to determine if a given pixel belongs to a class or not. Class signatures are derived by manually selecting small areas that are known to belong to a certain class. These areas are called training samples. In this way training samples are used for the determination of the characteristics of each class. The training samples in this research have been selected based on interpretation of the height data and by looking at independent reference data. In this case the palaeogeographic map was used mentioned in Section 3.3. In Figure 5.8 the location of the selected training samples are shown with different colours for the different classes. The classification results largely depend on the training samples taken by the operator. This is why several training sites have been taken for each class. In the software used for this research (ArcGIS) it was not possible to combine training samples from other tiles, which could have ruled out possible regional variations in class signatures. Using more tiles and training samples from different locations in the Alblasserwaard to derive class signatures could lead to improved classification results. This is also the case if the total area of the Alblasserwaard is divided in different regions and local class signatures are used for the classification process.

The training samples are used to calculate the mean, variance and covariance of all the established classes. Also a (Bayesian) probability function is calculated for the class data. Each pixel is then assigned to the class to which it most probably belongs. The method of the Maximum Likelihood Classifier has been implemented as one of the basic functions in ArcGIS.

Using the information from the bands mentioned above, it is possible to assign pixels to a certain class. In this research two cases were considered:

- **2 classes:** Consisting of the class 'Non-buried channel deposits' and the class 'Buried channel deposits'.



**Figure 5.8:** *The filtered height data, with the location of the training samples used for determining the class signatures. For the selection of the training samples additional data from the palaeogeographic map was also used.*

- **3 classes:** Consisting of 'Non-buried channel deposits', 'Buried channel deposits (top)' and 'Buried channel deposits (sides)'.

This distinction is made to determine the effect of using different amount of classes in the classification process. Buried channel deposits have different characteristics depending on the location across the channel. On top there is very little slope, the area is convex and relatively elevated compared to the surrounding area, the sides of the channel have a high amount of slope, are less convex and have less elevation difference with the surrounding area.

It is possible to take a look at the results of the classification process in chapter 6. However there are a number of additional steps needed to obtain these results. In the next paragraph it will be explained why the process of downsampling is applied and how this is implemented. Furthermore several additional automation processes were put into practice, these will also be described.

## 5.3 Downsampling

### 5.3.1 Moving window

The mentioned structural attributes describe the processing that leads to the calculation of the slope, curvature and TPI-value for a single pixel. For each pixel this process needs to be performed, so a

moving window is used. For the slope and curvature the window is squared and the size can be set to any odd number of pixels. For the TPI-value an annulus is used with an integer number of pixels for the distance of the internal and external diameter. The size of the moving window depends on the scale of the buried channels that we are looking for. This connotation of scale is called operational scale [8]. It is the scale at which a phenomenon operates; for example a city operates at a larger scale than an individual house. Different operational scales represent and carry different kinds of information. The effect of a small window will be that a local effect will have a bigger influence and small scale topography will be extracted, in contrast to a larger window where a more general trend will be calculated. So we must adjust the size of the window to match the operational scale of buried channels. If we take the characteristics of buried channel deposits into account (given in chapter 2), we know that their regular width is between 20 and 500 m. For this research an operational scale of 100 m was chosen. The pixel size is  $2 \times 2$  m, so this would translate to a window size of  $50 \times 50$  pixels. However, there needs to be a single centre pixel and this means that the window size needs to be an odd number of pixels. It was then chosen to use a window size of  $49 \times 49$  pixels for the calculation of the structural attributes slope, curvature and smoothed TPI. For the structural attribute TPI an annulus is used with an outside diameter of 49 pixels and an inside diameter of 39 pixels.

Other methods that could possibly be used (but are not explored in this thesis) to calculate the scale of buried channels to determine the optimal window size are: determining the scale factor [16] and the application of (semi)variograms to determine second order stationarity or to calculate the fractal dimension [11].

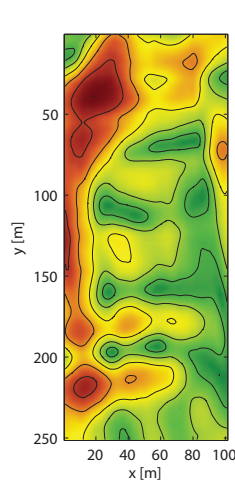
### 5.3.2 Point downsampling for plane fitting

The number of points used for calculating slopes depends on the size of the window that has been chosen. However, the computational load of the plane-fitting algorithm increases rapidly with an increasing window size. This is caused by the quadratic increase in the number of points associated with the window size. For example if a window with  $31 \times 31$  points is selected, there are 961 height values that can be used to determine the slope. Processing windows with many points is computationally intensive because for each slope calculation various matrix multiplications and one matrix inversion need to be performed. The computational complexity of matrix multiplication and matrix inversion is:  $O(n^3)$ . One of the ways to counter this problem is to reduce the amount of points in the window, this causes a reduction of the amount of calculations and thus decreases computation time. To reduce the amount of points a sub-group needs to be selected from the total group of points in the plane. This process is called downsampling. A uniform distribution of points throughout the window should be used to avoid the introduction of additional (directional) errors. Because the window is a rectangular grid, downsampling can be performed using systematic sampling [25], which selects points in every  $i^{th}$  row and  $i^{th}$  column of the window. The result is a uniform distribution of points.

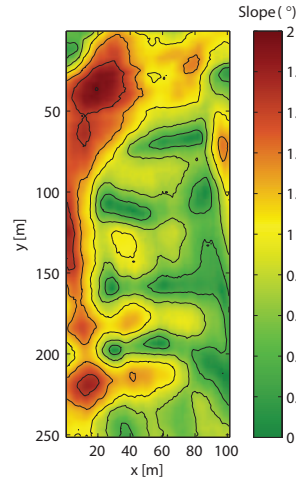
#### Errors

Using a reduced amount of points for the calculation of the slope reduces the reliability. To give an estimation of the quality of the planes calculated with a downsampled window, a representative area in the dataset has been selected for testing. This area contains buried channel deposits on the

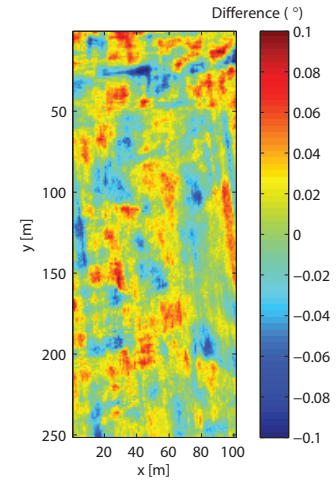
west side of the area covered by the dataset and is located at  $x = 118,000$   $y = 431,249$  in RD coordinates. A window size of  $49 \times 49$  m is used. The slopes calculated using downsampled data are compared to slopes that have been calculated with all available data. Figure 5.9 to Figure 5.11 gives an example of both cases and the small differences (or errors) that occur due to downsampling of the data. The differences have an average of around zero and do not deviate more than 5% compared to the reference area. Some systematic differences or patterns can be observed, which is probably due to major/rapid changes in height. Despite these differences both the results are very similar, this indicates that downsampling can be applied. What can now be investigated is the maximum amount of downsampling that is still acceptable.



**Figure 5.9:** Reference area where the slopes have been calculated using 100% of the points.



**Figure 5.10:** Reference area where the slopes have been calculated using 10% of the original number of points.

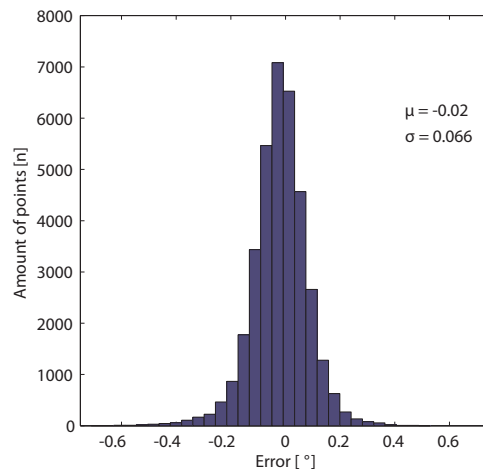


**Figure 5.11:** The difference between Figure 5.9 and Figure 5.10.

### 5.3.3 Downsampling optimization

All of the individual errors between the slope calculated using a full dataset and the slope calculated by the same dataset downsampled to 30% of the points are illustrated in Figure 5.12. Because the histogram approximately represents a normal distribution with a zero-mean, the standard deviation has been chosen as an indicator of the spread of the errors. To decrease processing time as much as possible we aim to minimize the number of points without significantly increasing the errors. To find a solution to this problem the previously mentioned process of downsampling has been repeated for varying downsampling rates. In Figure 5.13, the standard deviation of the error has been plotted against the percentage of the number of points used. When a majority (100 - 30%) of the points are used, a slow but steady rise of the standard deviation of the errors occurs. If only 30 - 10% of all points are used the standard deviation of the errors starts to rise rapidly, with a steep increase when less than 10% of the points are used.

Because errors introduced by downsampling could result in classification differences, the next step is to determine what errors are still acceptable. Using all the points available, the standard deviation of



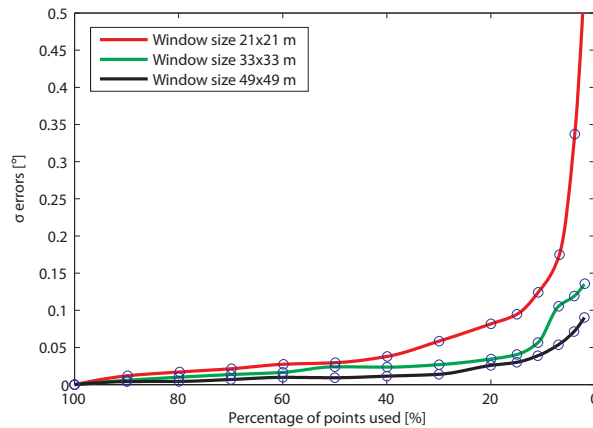
**Figure 5.12:** Histogram of slope errors between a reference area and a downsampled area with 30% of the amount of points. The window size used in this simulation is  $21 \times 21$  m.

the slope in the selected area is  $\sigma = 0.50^\circ$ . If a 10% increase of the standard deviation is taken as an acceptable amount, the additional standard deviation becomes:  $\Delta\sigma = 10\% \cdot 0.50^\circ = 0.05^\circ$ . In Figure 5.13 it can be seen where this value coincides with the percentage of points used for each of the window sizes. For a window size of  $21 \times 21$  m this is at 40% and for a window size of  $33 \times 33$  m this is at 15%. In the case of a larger window ( $49 \times 49$  m), downsampling to 7% of the total number of points in the plane can be used. However, errors introduced in this step will propagate into the final result, meaning that the downsampling percentage for slope determination should not be chosen too close to this number. This leads to the conclusion that downsampling of the number of points in a window can increase processing speed without significantly increasing the slope errors. The percentage of points used for slope calculation depends on the window size. For a  $49 \times 49$  m window this is at least 7%. Due to propagation of the errors, it is advisable to use a higher percentage, for instance 10%.

### 5.3.4 Downsampling for other attributes

The same downsampling principle could also be applied to the other structural attributes (curvature and TPI). The algorithm for the calculation of the curvature is relatively similar, so it is assumed that the same percentages can be used for downsampling. For calculating the TPI value no downsampling needs to be applied in practice, because this algorithm has a lower mathematical complexity.





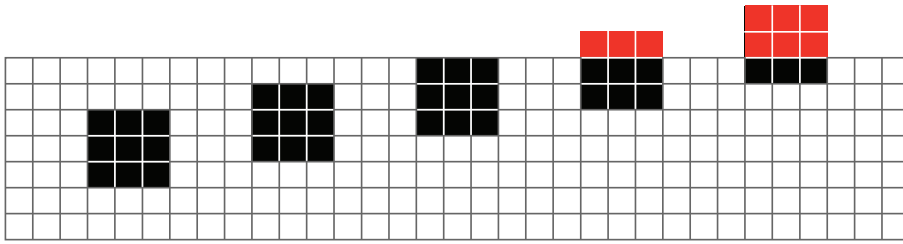
**Figure 5.13:** Standard Deviation of errors caused by different amounts of downsampling of points used to estimate planes. The colours represent the different window sizes used in this simulation.

## 5.4 Resolving Edge Effects

### 5.4.1 Moving window operations on tiled datasets

As mentioned in paragraph 3.1.2, the AHN-2 data has been converted into tiles of  $2,500 \times 2,000$  pixels. In practice these tiles make the data manageable because it simplifies data transfer, data modification and performing calculations on the data. The downside is that when a calculation on tiled data is done using a moving window, this window operation can not be calculated near the edges of the tile because there is no data available beyond the edge. This effect is illustrated in Figure 5.14. All four of the derived attributes which have been described in the beginning of this chapter are subject to this effect. A possible solution is to append all of the tiles together to form one very large tile to do the processing on. But this would lead to the exact situation that the tiles were supposed to solve, a very big dataset that is difficult to manage and to do processing on. Another option is to leave the tiles intact and perform the moving window operation up to the edges of that tile using only the data from that specific tile. This is not a desirable solution either, because elevations could significantly change just beyond the border of a tile. This information is then not used because the moving window can not access this information. Eventually this can lead to results that suddenly change at a border and do not connect across tiles.

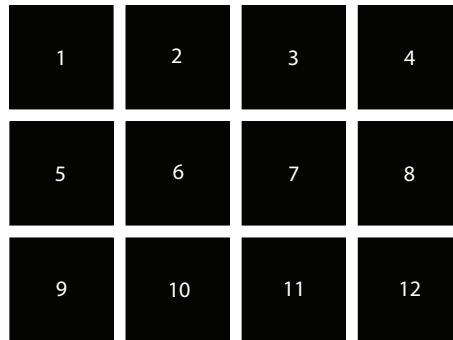
The window uses information from the surrounding pixels to calculate a new value for the centre pixel. If the window is near an edge information about pixels that lie beyond the edge is needed. The maximum distance of these pixels is half the window size (in horizontal and/or vertical direction). It can be stated that if a tile has size  $(p \times q)$  and the square moving window has an odd number of pixels  $(L)$ , only the data inside the tile can be processed up to pixel:



**Figure 5.14:** A moving window (black) at various stages of processing, moving upwards a pixel at a time. The red values represent that there is no data available.

$$\begin{aligned}
 (x_{\min}, y_{\min}) &: \quad \frac{1}{2}(L - 1), \quad \frac{1}{2}(L - 1) \\
 (x_{\max}, y_{\max}) &: \quad (p - \frac{1}{2}(L - 1)), \quad (q - \frac{1}{2}(L - 1))
 \end{aligned}
 \tag{5.14}$$

If many adjacent tiles are processed this leads to a kind of 'checkerboard-effect'; all the sides of the processed tiles have missing data at the edges because of the moving window. This effect can be seen in Figure 5.15



**Figure 5.15:** An example of multiple processed tiles where the moving window leads to a 'checkerboard-effect'.

The solution to this problem is to import the data from the edges of the neighbouring tiles and temporarily append it to the original tile. Using this 'oversized' tile the moving window calculation can be applied on the full extent of the original tile. The size of the appended data depends on the location of the tiles; horizontally adjacent tiles have a different size than vertically adjacent tiles. For pixels in the corner of the original tile, data needs to be imported from diagonal neighbours. The following three sizes for appending data are discerned:

$$\begin{aligned}
 \textit{Horizontal} &: \quad \frac{1}{2}(L - 1) \times q \\
 \textit{Vertical} &: \quad p \times \frac{1}{2}(L - 1) \\
 \textit{Corners} &: \quad \frac{1}{2}(L - 1) \times \frac{1}{2}(L - 1)
 \end{aligned}
 \tag{5.15}$$

After the appending of this data the new 'oversized' tile has size:

$$(p + L - 1) \times (q + L - 1)$$

The required processing can be performed on this new tile. It will be exactly large enough to let the moving window get a processed tile of size  $(p \times q)$ . After the processing has been completed the temporary appended data are discarded and the next tile is processed in the same manner.

### 5.4.2 Topology

In order to correctly construct the oversized tile, we have to determine which tiles are adjacent to each other. In other words: we have to determine all the neighbours of each tile. In the GIS-world this is known as topologic relationship [11]. In this research only the 'full neighbours' are extracted, which are the horizontal and vertical adjacent tiles. The set of these four tiles neighbouring a tile ( $t$ ) is denoted as  $N_4(t)$  and is called 4-adjacency. In Figure 5.16 an example is given of an area which has been divided into 12 tiles. The full neighbours of tile number 6 are:  $N_4(6) = t(2, 5, 7, 10)$ .

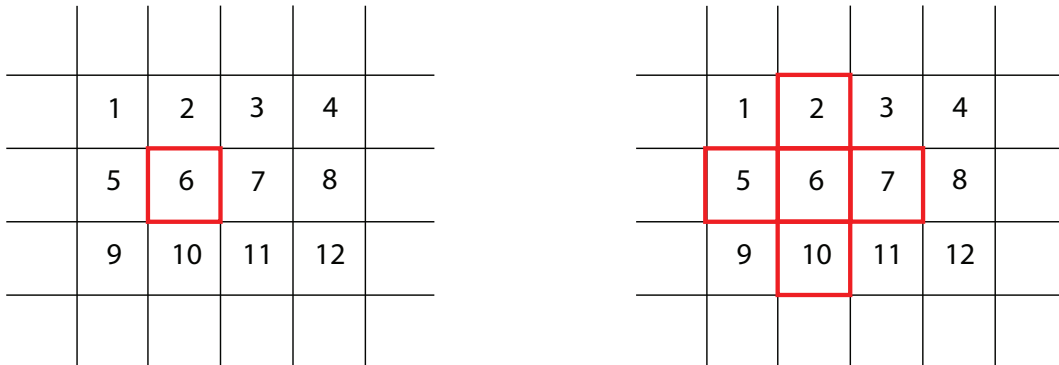


Figure 5.16: Basic topologic relationship - Full neighbours.

Using only the full neighbours introduces a possible error at the four corners of a tile, because at those locations the diagonal neighbours also influence the result. Compared to the area of the full neighbours, the area of the diagonal neighbours is relatively small and the introduced error would be negligible. This is the reason that we chose not to increase algorithm complexity to try to add the diagonal neighbours.

Determining the neighbours of each tile requires a number of steps: the coordinates of all four corners of the tiles must be derived from the meta-data of the file (coordinates of the top left corner, size of the tile  $(p \times q)$ , pixel size on the ground). If this is done for all tiles, the adjacent tiles can be extracted by determining which corner pixels are at a distance of one pixel from each other. Finally the output is written to a specified file so it can be used for further processing. The procedure used in this research has been adapted for use with different tile sizes. This is because in a few cases there is a difference in the size of the tiles in the FLI-MAP dataset. In some cases border pixels were pre-assigned to the wrong tile, causing these tiles to be one pixel wider or higher than others tiles. A limitation of this procedure is that a tile can only have one neighbour on each side. There is also no check if neighbouring tiles have overlap, but this effect should not occur if the data has been divided strictly.

## Chapter 6

# Results and Validation

The entire dataset was filtered according to the methods described in chapter 4, followed by the calculation of the four structural attributes and the classification process explained in chapter 5. In this chapter the calculated structural attributes are shown and the result of the classification is given. The result is filtered and interpolated to get a full coverage of the area. Finally, validation is performed based on two independent datasets.

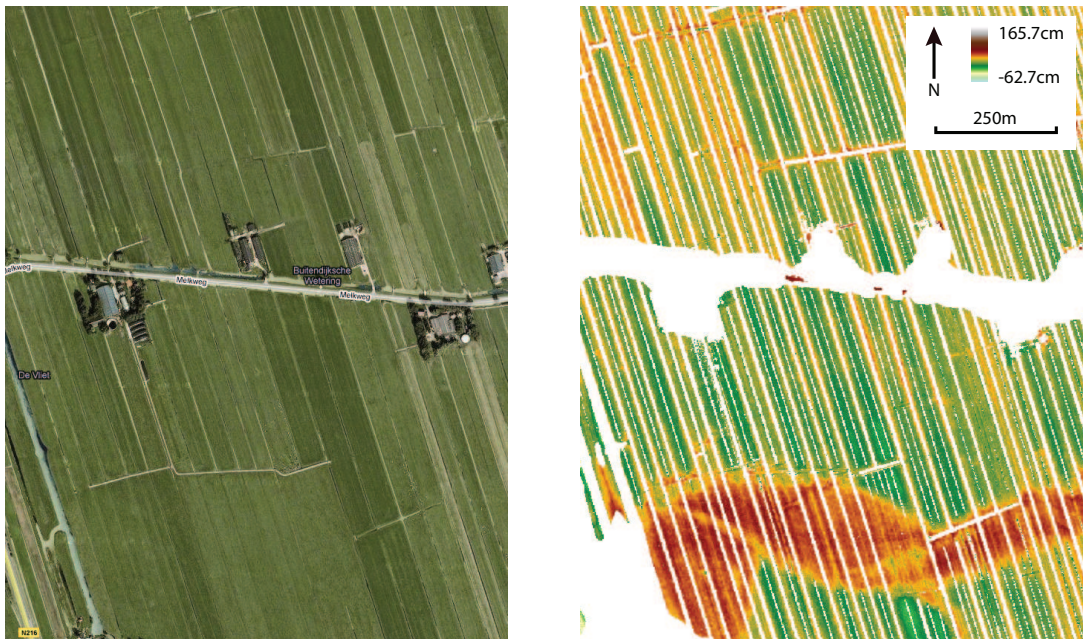
The full dataset covers an area of  $314 \text{ km}^2$ , but we are mostly interested in details at the 10 - 100 m level. This makes the (intermediate) results difficult to visualize, so a compromise between detail and overview was made. A single tile has been taken as an example and is shown throughout this chapter, an overview of the whole dataset is also shown. A different approach to the problem is presented in the form of an online GIS viewer.

### 6.1 Structural attributes

In Figure 6.1 the filtered height data of the single example tile is shown together with an aerial photograph of the same area. This tile is located in the centre of the Alblasserwaard. It can be seen that in the Southern part of the figure there is an elevated area probably caused by buried channel deposits. The large white area in the centre of the figure is an area that has been filtered out, it is a straight canal and road with multiple farms and other buildings on both sides. What can also be seen is the strong North-South parcel and drainage pattern with only a few perpendicular channels.

From the height data the four structural attributes were calculated. These are shown in Figure 6.2 a-d. If we analyze each structural attribute, there are several properties that can be observed:

- **Slope:** At the location of the buried channel deposits changes in slope can be seen at the surface. On both sides of the channel there is a gradient that ranges up to 4 degrees. On the top of the channel the surface is level. This pattern occurs along the flow-direction of the channel.

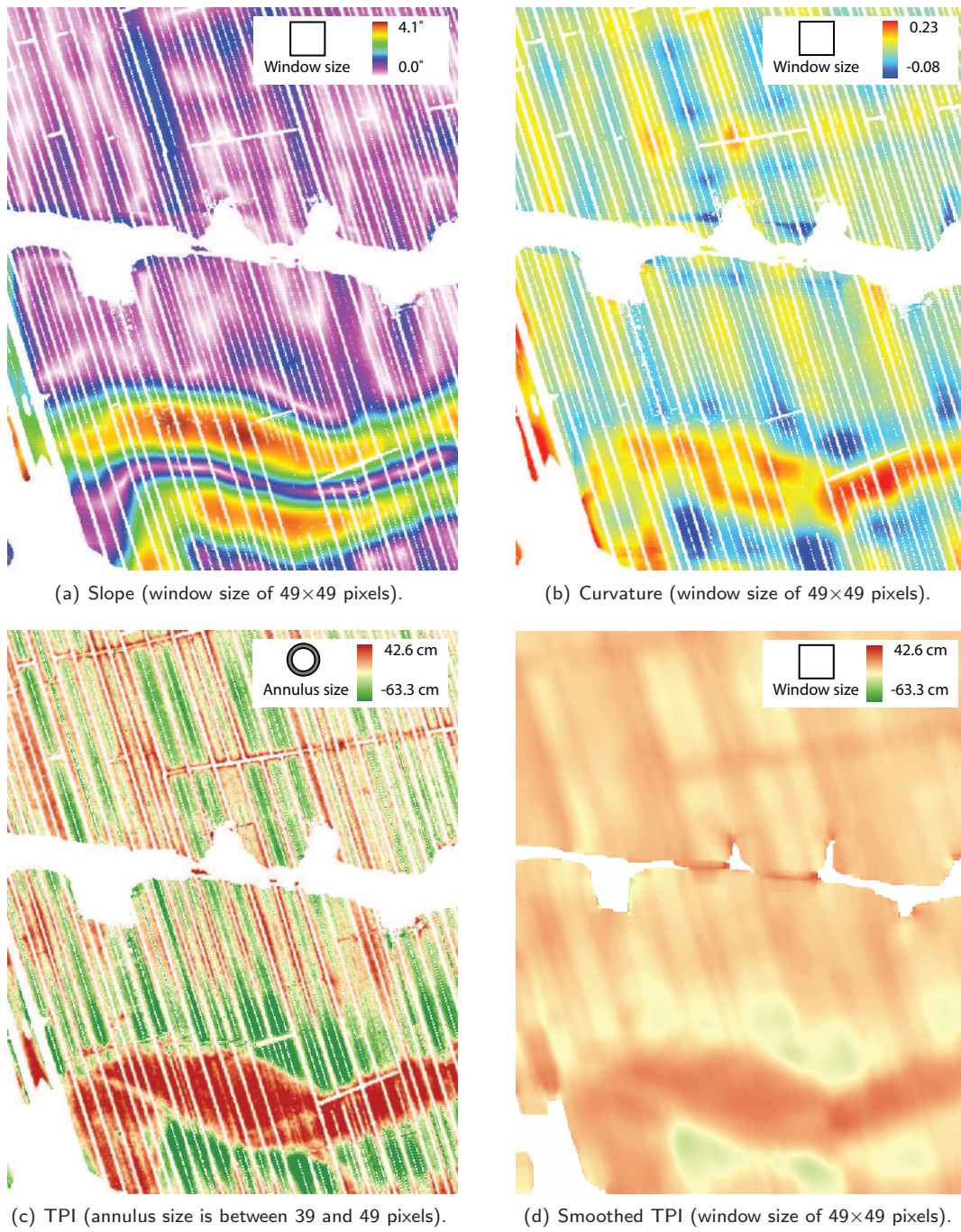


**Figure 6.1:** Aerial photograph and corresponding filtered height data. The location of the lower-left corner is in RD coordinates:  $x = 436,350$   $y = 120,000$ .

- **Curvature:** The buried channel deposits causes the surface to be convex, because it 'bulges' outwards from the surrounding landscape. But not only channel deposits have this property, also other areas show the same characteristics. For instance the road at the bottom left of the area is also convex.
- **TPI:** It can be clearly seen that buried channels deposits cause the surface area to be relatively higher than the surrounding area. Also some of the fields, especially the edges of fields, show this characteristic. It is caused by deposition of dredged material on the areas besides the canals. There is still a lot of detail in the image compared to the other structural attributes because the height difference between single height points and the average surroundings is calculated.
- **Smoothed TPI:** In this image there is less detail compared to the original TPI image. The amount of small objects that can cause classification errors has been reduced, leading to a result where larger topographic feature like the buried channel deposits are more prevalent.

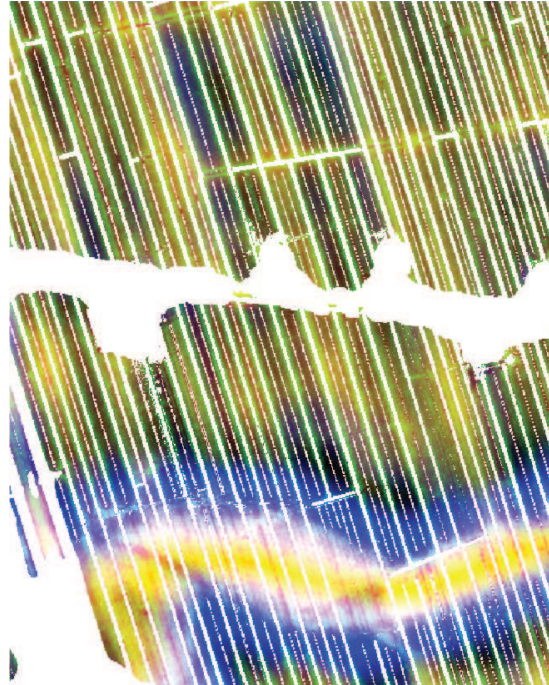
The four structural attributes are combined to form a single multi-band image. One of the ways to display such a multi-band image is to select three bands and display them as red, green and blue respectively, this is called a colour composite. Each combination of three attributes can be made and leads to a different colour composite. In Figure 6.3 the colour composite of the 'slope', 'curvature' and 'TPI' is shown, the 'smoothed TPI' is not shown. The three bands have only been selected for illustrative purposes and it does not influence the classification process because during classification the information from all four bands is used. The colours of the image in the Figure is a result of the selection of the bands and their associated colour, which have been assigned in an entirely arbitrary manner. The colours do not directly relate any more to the direct height but are a combination of





**Figure 6.2:** The four structural attributes calculated using the height data shown in Figure 6.1.

different properties. Further interpretation of the colour composite is difficult and because it is not used for any other purpose it is also not necessary.



**Figure 6.3:** *Colour composite using three of the four structural attributes. Slope in blue, curvature in red and TPI in green.*

## 6.2 Results of detection

All the multi-band images of the whole dataset have been classified using a Maximum Likelihood Classifier, which was described in paragraph 5.2.1. In Figure 6.4 the classified version of the example tile can be seen using different amounts of classes. If we first only look at the case where two classes are used, there are a number of things that stand out: there are areas in the lower-third of the image that are classified as buried channel deposits, this is what is expected if you look for instance at Figure 6.1. However not all of the buried channel deposits are detected, this is the case at the lower-left area and at the hole in the lower-centre. It indicates that there are still problems with the detection algorithms. The areas that are misclassified have properties that more resemble non-buried channel deposits. What might be the case is that because these areas are on top of broadest part of the channel, this causes them to have no gradient, and a low or negative curvature. This problem is related to the size of the windows used for the structural attributes: smaller windows result in the detection of smaller-scale phenomenon, larger windows result in larger-scale phenomenon. A balance had to be found between detecting small channels and larger channels so a window size was chosen based on the average size of the channels. This effect now causes errors in detecting some parts of the larger channels.

Besides the larger errors, the detection result also contains small errors. An example can be seen in the centre of Figure 6.4. In the area where the road has been filtered out there are a few small areas which have been classified as buried channel deposits. A solution to remove these areas will be explained in the next section.

The difference between using two or three classes is that there is an additional side part of the channel that is detected and there is a slight increase of the amount of buried channel deposits. Furthermore there are more false classification using three classes. This effect can not only be seen in the example tile, but is present throughout all of the other tiles. The additionally detected sides of the buried channel deposits do not contribute to the detection of new parts of the channel. When two classes are used there is a much cleaner result. This is why for the validation process only the result from the classification with two classes is used.

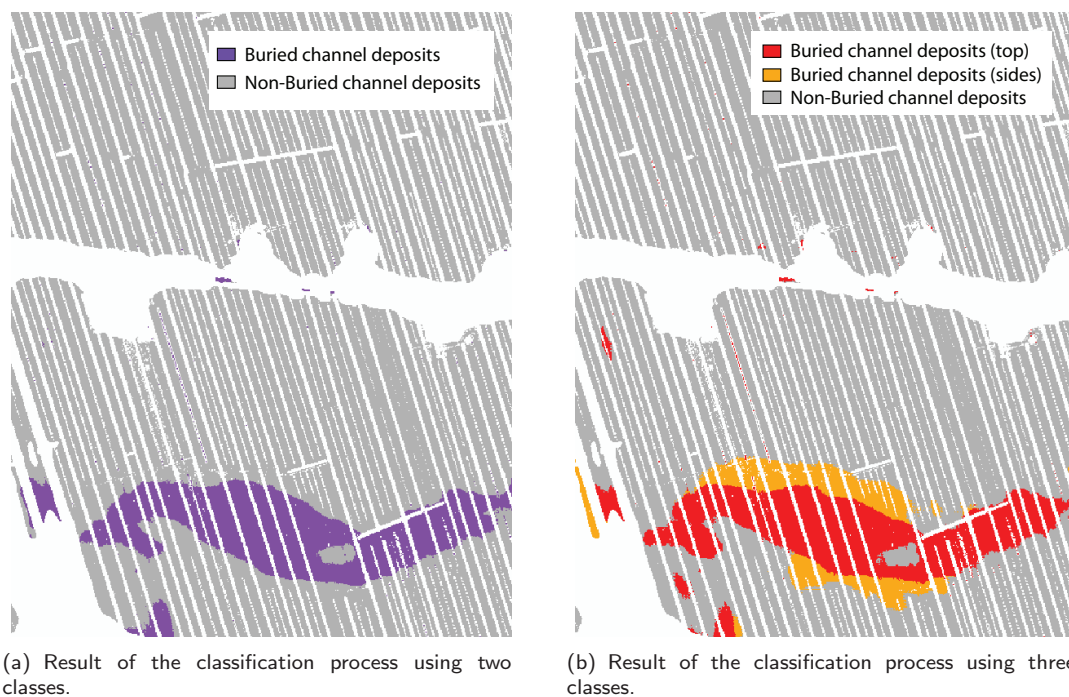


Figure 6.4: Result of the classification process.

### 6.3 Filtering and Interpolation

The detection results are filtered to remove small classification errors, followed by the interpolation of empty areas that were created by the filtering applied in chapter 4. How the process of filtering and interpolation is performed is described in the following section.



### 6.3.1 Filter

The morphological operator 'majority filter' [26] is applied to remove some of the misclassification and to smooth rough edges. In this operation the neighbouring pixels of a single pixel are analyzed and the value of this pixel is changed to the value which agrees with the majority of the neighbours. In this research the direct horizontal, vertical and diagonal neighbours ( $N_8$ ) are used, this resembles a  $3 \times 3$  window. An example can be seen in Figure 6.5.

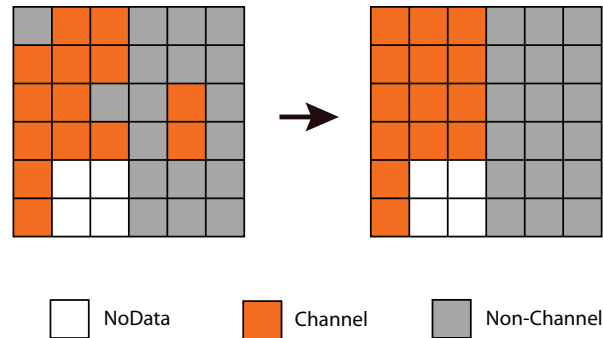


Figure 6.5: Example of the effect of a Majority Filter.

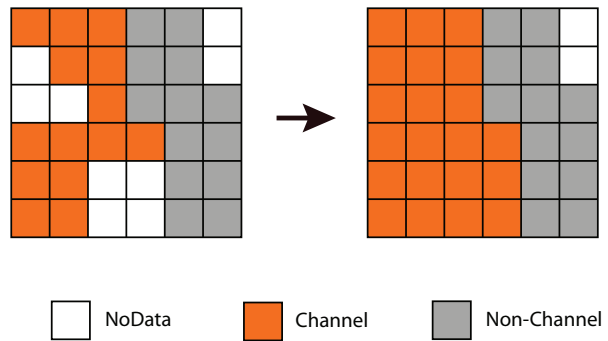
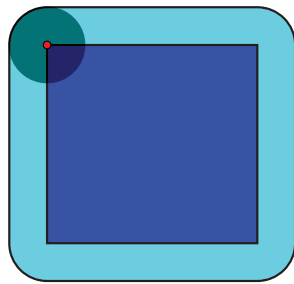
### 6.3.2 Conditional dilation

Dilation is a morphological operator [31], it uses a structuring element to expand a given shape. The structuring element used in this research is a  $3 \times 3$  window. This shape has been chosen because of the straightforward implementation in a matrix environment, the size is based on the size of the areas that needed to be filled in (like canals and small roads). The dilation operation was only performed on empty pixels, so the already existing classification would not be affected. In this way the dilation operation becomes an interpolation technique. By imposing a condition on the dilation, it becomes a conditional dilation [15]. The conditional dilation of shape A by structuring element B, with conditional shape C is defined by:

$$A \oplus_C B = \bigcup_{a \in A} B_a \cap C \quad (6.1)$$

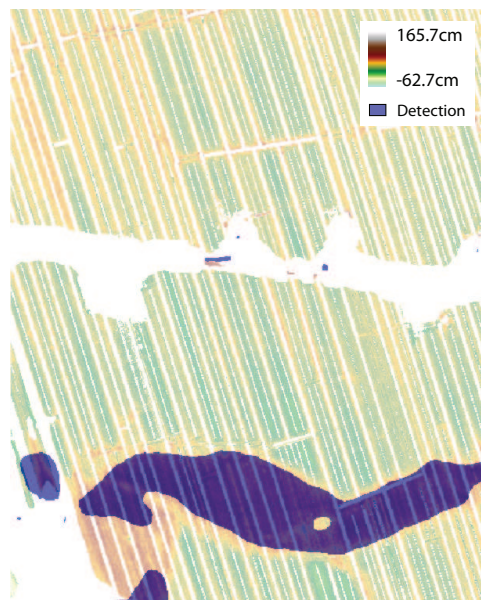
In our case shape A is the area of the class 'Buried channel deposits', structuring element B is the  $3 \times 3$  window and the conditional shape C is the area that contains the empty pixels. An example of normal dilation is seen in Figure 6.6, an example of conditional dilation is seen in Figure 6.7.

This resulted in the closing of nearly all the gaps in the 'Channel' class, without enlarging the original detection result at the cost of the 'Non-Channel' class. If this method is applied after the majority filtering explained previously, the final classification result is smoother and uninterrupted, see Figure 6.8.



**Figure 6.6:** The dilation of the dark-blue square by a disk, resulting in the light-blue square with rounded corners (Figure by [20]).

**Figure 6.7:** The result of a conditional dilation of the 'Channel' class using a  $3 \times 3$  window as structuring element, with the condition that dilation can only take place on the 'NoData' class.



**Figure 6.8:** Small area of the final classification result after filtering and interpolation, in the background is the original height data. Compare with Figure 6.4(a).

In Figure 6.9 the final classification result of the Alblasserwaard is shown after the filtering and interpolation operations.

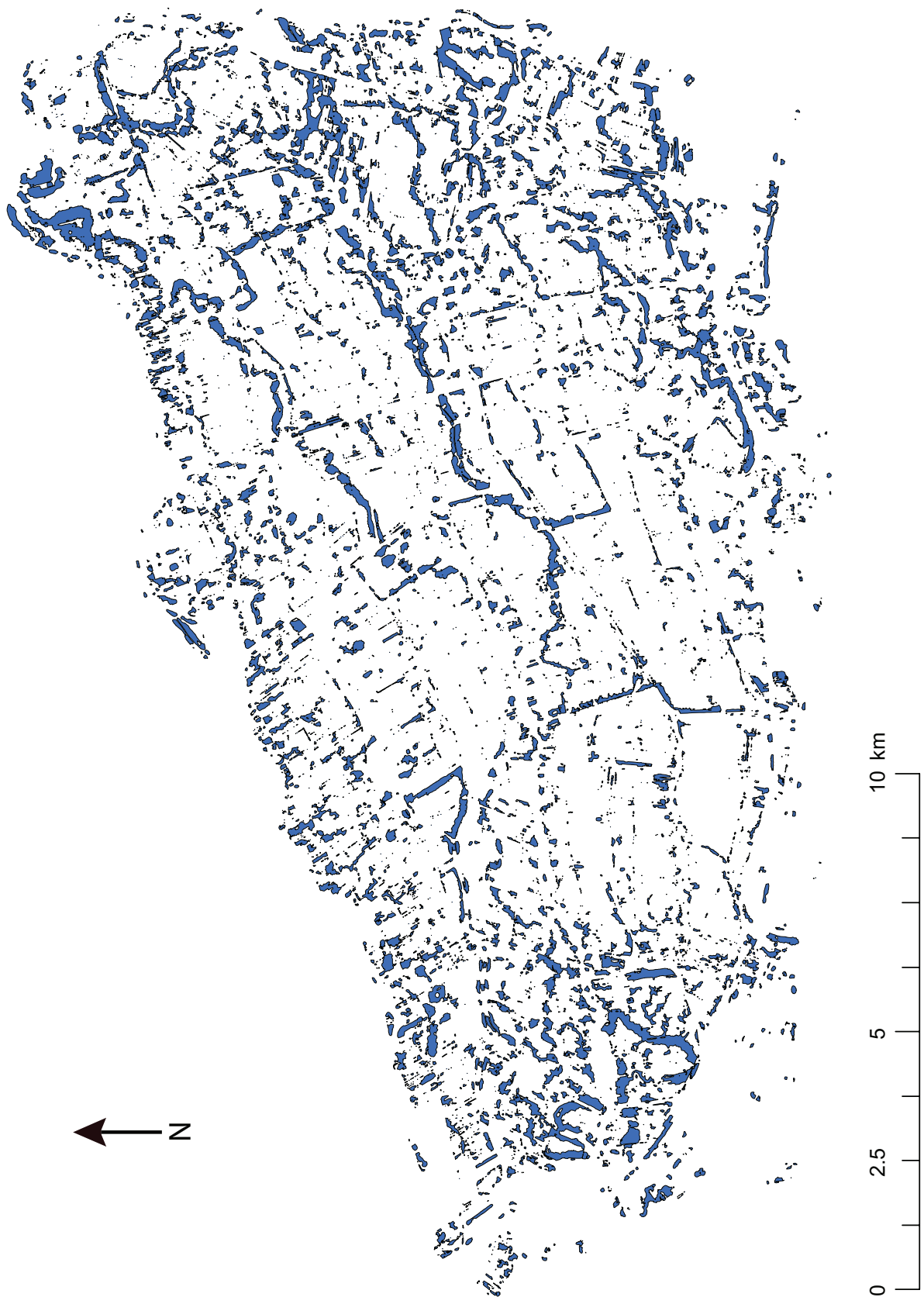


Figure 6.9: Final classification result of the Alblasserwaard after filtering and interpolation.

## 6.4 Validation

A way to assess the accuracy of the detected channel deposits is to compare it with a non-related dataset containing ground truth of the area. In this research validation is performed by comparing the detection result to the palaeogeographic map treated in chapter 3.3. This map is compared using two different methods: a numerical comparison and a visual interpretation of the differences. The detection result is also validated against another independent dataset, the shallow drilling measurements from the DINOloket.

### 6.4.1 Comparison with palaeogeographic map

#### Numerical comparison

To get an unbiased comparison (not based on human interpretation) between the detection result and the palaeogeographic map, a numerical comparison is performed. This is done by comparing the number of classified pixels in both maps to each other. The detection result is already in raster format, but the palaeogeographic map in vector format. To compare both maps on a pixel level the palaeogeographic map needed to be converted from vector to raster. After this was done, each pixel from one map was compared to the corresponding pixel in the other map. A pixel in one map may be the class 'Buried channel', but in the other map it may have been classified 'Non-Buried channel'. With two classes in each dataset, this results in four different options. A summation of all pixels in the different options was made and the results are shown in table 6.1.

This matrix is called the confusion matrix. On the diagonal entries the amount of corresponding pixels is shown. The two non-diagonal entries are pixels that do not correspond. The upper-right corner shows the amount of pixels that are Non-Channel pixels in the reference map, but classified as Channel pixels in the detection result. This is called the *commission error*, incorrectly including pixels to a class that belong to another class. The lower-left corner shows the amount of pixels that are Non-Channel pixels in the detection result, but are classified as Channel pixels in the reference map. This is called the *omission error*, pixels that should have been identified as belonging to a particular class but have been classified differently.

**Table 6.1:** *Confusion matrix of channel classification (numbers are the amount of pixels).*

		Reference map	
		Channel	Non-Channel
Detection result	Channel	2,777,458	4,813,182
	Non-Channel	18,439,011	52,375,000

Total number of pixels in this area is 78,404,651

The overall classification accuracy is the sum of the diagonal elements ( $d_{ii}$ ) divided by the total number of pixels ( $N_{tot}$ ):

$$\frac{\sum d_{ii}}{N_{tot}} = \frac{2,777,458 + 52,375,000}{78,404,651} \cdot 100\% = 70.34\% \quad (6.2)$$

This measure does not take chance-agreement into account. For example, if two totally random maps with only two classes are compared, half of the detected pixels will be assigned to the correct class by chance. This would lead to an overall classification accuracy of 50%. To correct for this effect another measure is needed. In this research Cohen's kappa coefficient ( $\hat{\kappa}$ ) is used, which is a statistical measure for the agreement of qualitative (categorical) items [13]. The kappa value indicates how accurate the classification output is after the random portion has been accounted for. The kappa coefficient ranges from -1 to 1, where  $\hat{\kappa} = -1$  is total disagreement and  $\hat{\kappa} = 1$  is total agreement. When  $\hat{\kappa} = 0$  this means that there is no relationship between the maps.

The kappa coefficient can be determined by applying Formula 6.3. Here  $N_{tot}$  is the total number of pixels,  $d_{ii}$  are the main diagonal elements,  $d_{i+}$  is the sum of the row and  $d_{j+}$  is the sum of the column.

$$\hat{\kappa} = \frac{N_{tot} \cdot \sum d_{ii} - \sum (d_{i+} \cdot d_{j+})}{N_{tot}^2 - \sum (d_{i+} \cdot d_{j+})} \quad (6.3)$$

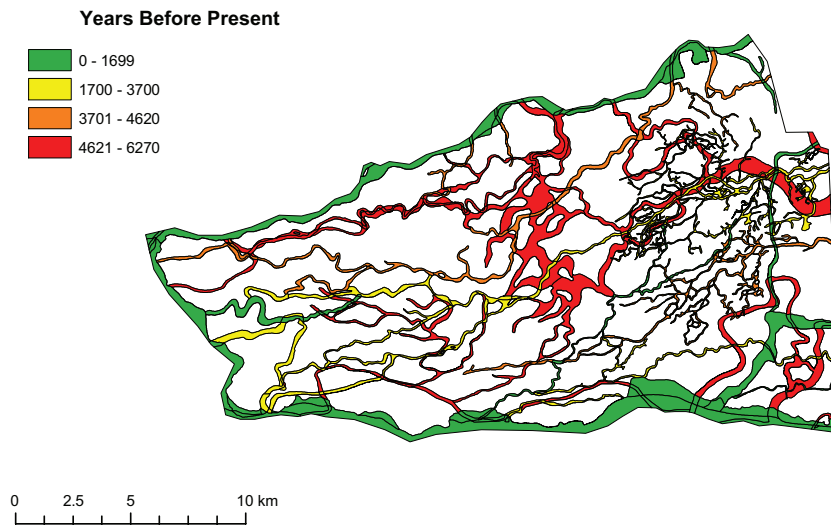
Where in this case:  $\sum (d_{i+} \cdot d_{j+}) = (2,777,458 + 4,813,182) \cdot (2,777,458 + 18,439,011) + (18,439,011 + 52,375,000) \cdot (4,813,182 + 52,375,000)$

If we calculate kappa for the detection result and the palaeogeographic map it leads to:  $\hat{\kappa} = 0.06$ . This number translates to a slight agreement [22]. It is caused by the fact that there are large areas where the palaeogeographic map indicates that there have been ancient rivers, but these areas do not match the detection map. This corresponds with the relatively large omission error (amount of pixels in the lower-left corner of Table 6.1). There are also a lot of smaller areas where buried channel-deposits have been detected, but are not indicated on the palaeogeographic map.

An hypothesis is that only buried channel deposits from certain time periods have lead to landscape inversion. So only these channels can be detected using the previously explained method. This theory is further investigated in the next section.

### Digital comparison using time periods

The approximate age of the different channels in the palaeogeographic map can be used to investigate whether there are certain time periods where buried channel deposits can, or can not, be detected. The channels are separated into four different periods based on the time the channels were abandoned. These periods have been manually chosen based on the distribution and the amount of channels that were abandoned in these periods. The channels from different periods have been visualized in Figure 6.10.



**Figure 6.10:** Palaeogeographic map of the Alblasserwaard with different selected periods of channel abandonment.

For each of the four periods we again want to calculate the amount of agreement with the palaeogeographic map. For this the kappa coefficient is not suitable because we do not have two maps with the same features that can be compared. The palaeogeographic map can be split into 4 separate maps with different time periods, but the detection result is a combination of all the channel deposits and can not be split up. The solution applied in this case is to calculate the relative correct classification. This is the amount correctly classified pixels divided by the sum of the omission error and the amount of correctly classified pixels (see equation 6.4).

$$\text{Relative correct classification} = \frac{\text{Correctly classified}}{\text{Omission error} + \text{Correctly classified}} \cdot 100\% \quad (6.4)$$

The results are shown in table 6.2.

**Table 6.2:** Matrix of correct channel classification per period

Abandonment period (yBP)	0-1699	1700-3700	3701-4620	4621-6270
Channel pixels (correctly classified)	544,341	930,628	664,894	637,089
Non-Channel pixels (omission error)	7,587,222	2,514,363	1,720,848	6,616,775
Relative correct classification (%)	6.7%	27.0%	27.9%	8.8%

What can be concluded is that there is a relation between the relative correct classification and the different periods. There are two periods that have a low classification rate, and two periods that have a higher classification rate.

If we now only select the channels from the two middle periods (1700 - 3700 yBP and 3701 - 4620

yBP) and again calculate the confusion matrix, this results are shown in table 6.3.

**Table 6.3:** *Confusion matrix of channel classification with channels from selected periods*

	Reference map	
	Channel	Non-Channel
Channel	1,594,995	5,994,565
Non-Channel	4,235,750	66,579,157

The total number of pixels in Table 6.3 is: 78,404,467. That is a difference of 184 pixels compared to the previous calculation (Table 6.1). These differences can be accounted for due to rounding errors when the vector-files are converted to raster-files.

By applying equation 6.2 it is calculated that the overall classification accuracy is 87.0%. Using equation 6.3 the non-random correspondence between the detection map and the channels from the selected periods in the palaeogeographic map result in:  $\hat{\kappa} = 0.21$ . This is a fair agreement [22], but still quite low for two maps that should indicate the same feature. Why this is will be investigated in the next section, where another comparison is made based on visual comparison.

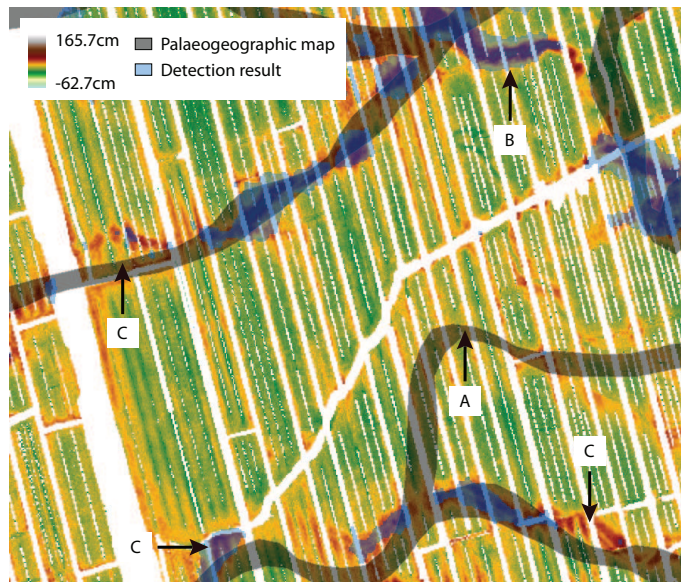
### Visual comparison

There are various differences between the palaeogeographic map and the detection results, based on a visual interpretation these differences can be grouped. A summary of what types of differences occur is given below.

- **Channels not detected in FLI-MAP data:** As already determined in section 6.4.1, some channels can not be detected using height data. An example can be seen at location A in Figure 6.11.
- **Channels partly detected in FLI-MAP data:** In many cases only parts of the channel are detected, in these cases the signal of buried channel deposits are not detectable anymore. There are many possible reasons for this (described in chapter 4.4.1), for instance: plowing, leveling, vegetation, buildings etc. Partial detection of buried channels is a problem caused by the fact that automatic interpretation is still inferior to human interpretation.
- **Missed channels in palaeogeographic map:** There are instances where buried channel deposits can be clearly seen and detected in the height data, but are not shown in the palaeogeographic map. An example can be seen at location B in Figure 6.11.
- **Incorrect channel locations:** In approximately 50% of the cases there is some form of offset between the detected channel and the palaeogeographic map. In most cases it is a missed bend in the river, but in other cases the channel runs parallel. This could be due to a wrong interpretation of drilling data or the resolution of sampling (drillings) was not high enough. Multiple examples are shown at locations C in Figure 6.11.
- **Incorrect channel widths:** A common difference between the classification results and the palaeogeographic map is the channel width. Compared to the channels in the palaeogeographic



map the channels in the detection result are sometimes significantly wider (more than 150%) or significantly smaller (less than 50%). It is not exactly clear why these differences occur, it could be an error in the palaeogeographic map or the fact that not all of the buried channel deposits result in landscape inversion and are therefore not detectable. More research should be conducted to find out what causes these differences.



**Figure 6.11:** Height data with the location of buried channel deposits from the palaeogeographic map and the detection result. Some of the differences can be observed. The location of the lower-left corner is in RD coordinates:  $x = 123,010$   $y = 435,410$ .

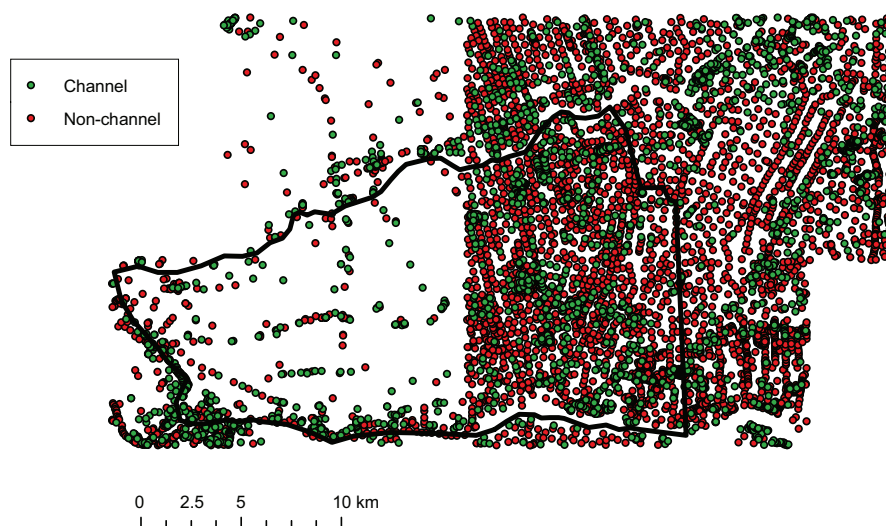
#### 6.4.2 Comparison with DINO Data

The drilling data from the DINOloket has been described in chapter 3.4. All the 6673 individual drillings were imported automatically and information like name, coordinates and depth of individual layers was extracted. An example of a drilling is given in Appendix C. Data from a larger region than the Alblasserwaard was supplied, so the 2,680 drillings that are inside the study area of the Alblasserwaard were selected.

After importing and extracting the DINO data, each drilling was automatically analyzed to determine if the drilling was part of a buried channel deposit. This was done by applying a basic filter: search for a sand layers that are cumulatively more than 3 meters in thickness in the shallow subsurface (between 3 and 12 meters). If more than 3 meters of sand was found, the drilling was classified as 'buried channel deposits'. In all other cases the drilling was classified as 'Non-buried channel deposits'. These numbers were chosen based on the characteristics of buried channel deposits, as explained in chapter 2. The reason why the top 3 meters are discarded is that sand layers can be present there due to other reasons, like construction works. It is also very unlikely that sand layers so close to the surface were deposited due to ancient river channels. The analysis of all of the drillings is shown in Figure 6.12. However, automatic interpretation of drillings is prone to errors. Only extensive human



analysis using the surrounding drillings can lead to the final interpretation of a single drilling. This means the amount of correctly interpreted drillings is unknown, but what can be stated is that it gives an indication.



**Figure 6.12:** Analysis of all the supplied drillings of the DINOLoket, the black line indicates the contour of the Alblasserwaard.

If we compare the drillings with the detection result and the palaeogeographic map we get the results shown in table 6.4. Both the detection result and the palaeogeographic map have a comparable amount of correct and incorrect classifications (sum of each row). They do however differ on the type of misclassification: in the detection result 719 of the 'Non-Channel' points should have been detected as 'Channel' points and 195 'Channel' points have been detected as 'Non-Channel' points. In the palaeogeographic map there are 486 'Non-Channel' points that have been detected as 'Channel' points and 499 'Channel' points that have been detected as 'Non-Channel' points, this is nearly equal. A possible reason is that the palaeogeographic map covers a larger absolute region than the detection result, this results in more 'Channel' points compared to the detection result.

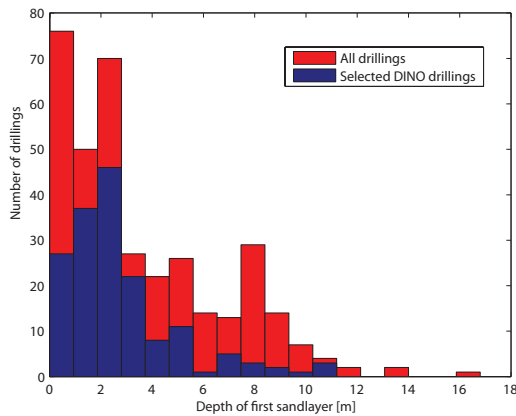
**Table 6.4:** Comparison between the drilling data from the DINOLoket, the detection result and the palaeogeographic map.

	Detection result		palaeogeographic map		
	Channel	Non-Channel	Channel	Non-Channel	
<b>DINO data</b>	Channel	169	719	402	486
	Non-Channel	195	1,597	499	1,293

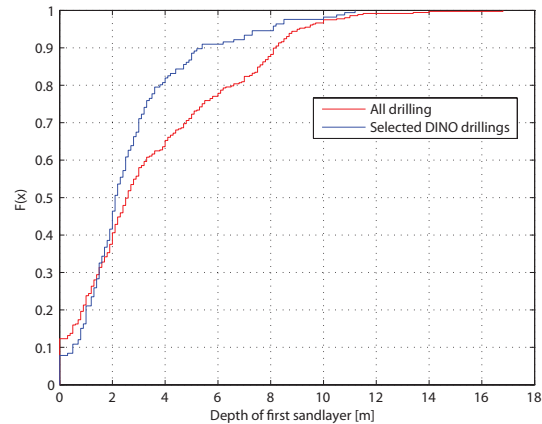
### Depth of the detected buried channel deposits

Possible users of the detection results map are not only interested in the age of the detected channels, but also in the depth of these channels. Depth information can be acquired by relating the location of detected channels with the detected sand layers in the DINO-data. The histogram of Figure 6.13 gives an indication of the depth of the shallowest sand layers at the location of the detected buried channel deposits. The red bars indicate that all the available drillings have been used, the blue bars indicate that only drillings classified as 'buried channel deposits' have been used (using the selection process described previously in this paragraph). What can be seen in the Figure is that sand layers have been detected in all drillings at the location of the detected buried channels deposits. What can also be seen is that for nearly all the drillings, the most shallow sand layers are within the first 10 meters. The blue bars indicate that there is at least an accumulated 3 meters of sand layers between 3 and 12 meters depth. The difference between the blue and the red bars can possibly be explained in two ways: there are locations where there are relatively thin sand layers (less than 3 meters in total) or it could be due to the measurement method. Most of the shallow drillings have been made manually with a gouge (in Dutch: "guts"). This equipment can not penetrate through sand very well, so there is sometimes no data available on the subsoil underneath a sand layer of several meters. This also means that it is not possible to give a full picture of the thickness of the sand layers in relation to the location of the detected buried channel deposits.

Using the data from the histogram a cumulative distribution function can be made, as seen in Figure 6.14. From this figure it is possible to derive the probability that, in an area of the detection result, a sand layer will be present in the subsoil up to a certain depth. For example, the probability that a sand layer is present in the first 2 meters is approximately:  $F(2) = 0,4$ .



**Figure 6.13:** Histogram of the depth of the shallowest sand layers in the DINO-data at the location of the detected buried channel deposits.



**Figure 6.14:** Cumulative Distribution Function of the depth of the shallowest sand layers in the DINO-data at the location of the detected buried channel deposits.

It can be seen that all of the sand layers in the detection result have a depth of less than 10 meters. But from this result it can not be concluded that this landscape inversion is caused by the most shallow sand layers. It could be that there are multiple layers on top of each other and the detection result relates to a deeper layer. It does seem more likely that the shallower sand layers are detected because

they have been deposited more recently and are therefore more prone to the effect of differential compaction. Different datasets and methods are needed to say more about this subject.

## 6.5 Final product

One of the research objectives was to create a geo-referenced channel map product and to investigate what the best way is to visualize and distribute/store the data.

The final product can be used as a tool for the detection of buried channel deposits within a time frame (4620 - 1700 yBP) and the related depth (up to 10 meters). A relation can be given to determine what the chance is that a sand layer will be present up to a certain depth at the location of the detection result. However the thickness of this sand layer can not be given. Because of the uncertainties in the detection and the differences found during validation, it does remain important that field validation is performed. Therefore the practical use of the detection map is low, although the map could be used for updating the current palaeogeographic map and for further research.

If all of the DINO drillings are combined with the detection result, it may be possible to determine how deep the underlying sand layer is for each area in the detection result. Another product that could be developed is a probability map, indicating what the chance is that a sand layer is at a certain depth. Such a map could be useful for construction purposes, for example to determine how long foundation piles (in Dutch: "heipalen") need to be. Such map products could be the subject of a separate research project.

### 6.5.1 Visualisation

The final detection product is stored in vector format. This is far more efficient than a raster format and it saves a lot of storage space. Files can be stored locally on a computer for viewing and editing purposes or more complex tasks. It can be added as a separate layer in a GIS. This type of software contains all the necessary features to let users navigate through the data. A disadvantage is that the software needs to be installed locally creating a barrier for some users. The original FLI-MAP elevation raster dataset is very large (3 GB), which makes it difficult to work with. The only solution is to have this data stored locally.

A solution for the fact that not everyone has a GIS available is presented in the form of an online GIS viewer. This type of viewer makes it possible to visualize geographical data on every computer with the data stored on internet. Processing is performed on the server using a web map service, the results are converted to a picture and sent to the viewer of the user. The online Flex viewer used in this research is based on the free Adobe Flex framework, rendered by Flash Player. A major advantage of this type of viewer is that it can be fully adapted to fit the user.

A web map service and online viewer have been implemented with data used in this research project. How the viewer looks can be seen in Figure 6.15. A live demonstration can be found on: [www.possel.nl](http://www.possel.nl)

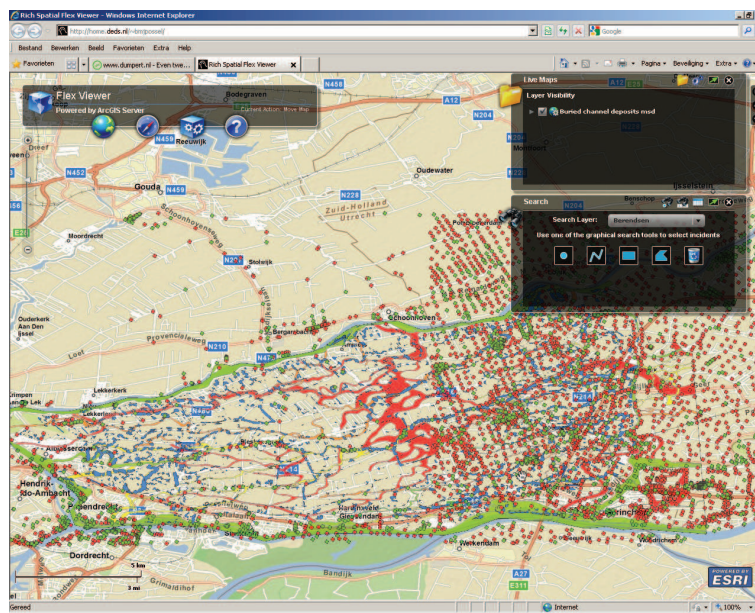


Figure 6.15: A demonstration of the online Flex Viewer.

## 6.5.2 Toolbox

A toolbox was created in ArcGIS so all the steps that were performed to create the geo-referenced channel map product can be repeated. The toolbox can be seen in Figure 6.16. The toolbox can be downloaded at: [www.possel.nl](http://www.possel.nl)

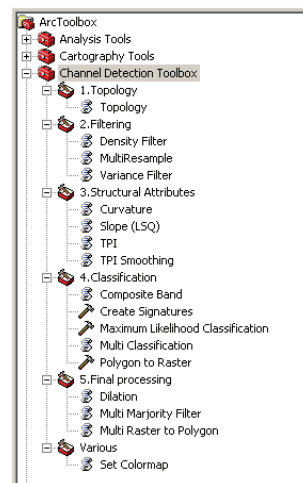


Figure 6.16: Overview of the Toolbox in ArcGIS.



## Chapter 7

# Conclusions and Recommendations

This thesis focused on the automatic detection of buried channel deposits using high resolution laser altimetry data (FLI-MAP). A method has been developed and demonstrated, followed by validation and visualisation of the results. In this chapter the conclusions are presented, treated according to the order of the research objectives that were defined in the introduction. Subsequently recommendations for further research are discussed.

### 7.1 Conclusions

#### 7.1.1 General

Buried channel deposits can only be detected at the surface in areas where topographic inversion has taken place. This limits the use of the detection method to these areas. Additionally, the method can only be applied in areas which have been (relatively) undisturbed by humans, these are predominantly farming fields that have not been leveled.

The resolution of the rasterized AHN-2 elevation dataset is more than sufficient to be used for detecting buried channel deposits. The data were resampled to  $2 \times 2$  m (the original resolution was  $0.5 \times 0.5$  m), to decrease processing time without compromising accuracy. This can be done because of the relatively large scale of the buried channel deposits in comparison to the pixel size.

AHN-2 data has advantages over AHN-1 data: a higher raw point density and higher accuracy in combination with better filtering techniques result an increased raster resolution (AHN-1 has a maximum resolution of  $5 \times 5$  m) and in a more reliable surface height determination. With AHN-1 the large buried channel deposits would probably be detected, but not the smaller channels. This leads to the recommendation to use resampled AHN-2 data or elevation data with the same specifications for the detection of buried channel deposits.

### 7.1.2 Filtering

The height dataset needs to be filtered to remove unwanted objects. These objects would otherwise lead to errors in the detection results. After several processing steps, it is possible to use the GBKN as a mask to remove topographic object. Additional filtering of the height data can be applied based on local variance and point density to further remove unwanted objects or areas.

The amount of false positives in the detection result decreases when a more stringent filter is applied beforehand. Because of the large amount of data it is advisable to use such a stringent filter even if this leads to the removal of usable data. This is better than applying a less stringent filter that leaves unwanted points in the dataset resulting in errors in the detection result.

### 7.1.3 Detection

To use as much of the available information contained in the height dataset in the classification process, various structural attributes were calculated. The attributes slope, curvature, TPI and smoothed TPI were combined to form a multi-band image. Using a Maximum Likelihood classification the multi-band image was classified in preselected classes.

In this research, systematic downsampling to 10% of the amount of points in a window is performed for the structural attributes. It is shown that this amount of downsampling increases computation speed without significantly increasing the errors.

There are many different local situations where the detection algorithm does not deliver optimal results. In many ways the problem can be seen as pattern recognition. In this case this process is difficult for a computer, but more obvious for humans.

### 7.1.4 Validation

If the detection result is compared to the palaeogeographic map, this leads to  $\kappa = 0.06$ . This indicates that there is only a slight agreement between these two maps. Many of the differences can be explained:

- Due to many changes in the landscape during the last centuries (construction of roads, buildings, etc.) the signal of buried channel deposits at the surface has been lost in many places. Therefore channels can not be automatically detected anymore at these locations.
- Despite the filtering of the height data there are still many objects and areas present that result in incorrect classifications.
- Not all buried channel deposits are detected, a part of this effect is related to the size of the windows used for the structural attributes: smaller windows result in the detection of smaller-scale phenomenon, larger windows result in the detection of larger-scale phenomenon. A balance had to be found between detecting small channels and larger channels, this now causes errors in detecting some parts of the larger channels.

The different channels in the palaeogeographic map were divided into four (manually selected) time periods and compared to the detection result. If the detection result is compared with the channels between  $\pm 1700 - 4620$  years Before Present from the palaeogeographic map this results in a slight resemblance with  $\kappa = 0.21$ .

### 7.1.5 Final product

The final product can be used as a tool for the detection of buried channel deposits within a time frame (4620 - 1700 years Before Present) with a related depth of up to 10 meters.

A relation can be given to determine what the chance is that a sand layer will be present up to a certain depth at the location of the detection result. However an indication of the thickness of this sand layer can not be given.

Because of the uncertainties in the detection and the differences found during validation, it is important that field validation is performed. Therefore the practical use of the detection map is low, although the map could be used for updating the palaeogeographic map and for further research.

The large amount of data and the large extent makes visualization difficult. GIS packages can solve this problem for individual users with the data stored locally. Web based GIS viewers make it possible to visualize the data on every computer for every user with the data stored on internet.

## 7.2 Recommendations

The TOP10NL [38] could be used for filtering instead of the GBKN. This dataset was not available at the time when the research for this thesis was performed. The advantage of TOP10NL is that it is object oriented, uses a vector-format and has a closed surface objects. This would make the processing work performed on the GBKN in chapter 4 obsolete.

There were 4 structural attributes used in the detection process. Research could be done to determine if other structural attributes could be included. For example the aspect of slopes. One or more of the (already available) structural attributes could be calculated a number of times using different window sizes. This could be done to obtain information about features that have different operational scales, for instance a large buried channel with a smaller side channel. A further study of the implication and (dis)advantages of this method could be conducted.

The window size for calculating the structural attributes was based on the operational scale which was again based on the average size of buried channel deposits. Other methods that could possibly be used (but are not explored in this thesis) to calculate the scale of buried channels to determine the optimal window size are: determining the scale factor [16] and the application of (semi)variograms to determine second order stationarity or to calculate the fractal dimension [11].

The classification method used in this research was Maximum Likelihood Classification. This was the only viable option available in the ArcGIS environment. It is highly recommended that more research is performed into the use of a different classification method.



Training samples from only a single tile were used to calculate the class signatures, this was due to the software used. However, regional differences in class signatures throughout the Alblasserwaard caused regional differences in the detection results. A first solution to this problem would be to use training samples from various regions throughout the Alblasserwaard area. But this would only reduce possible regional variations in class signatures. A second solution would be to divide the area into separate regions and classifying these areas independently using local training samples.

Only a rude comparison was made between using two and three different classes of training samples for the detection. The amount of classes could be further investigated, because it is not clear if other classes exist.

Other datasets from different sources could be integrated into the detection process, for instance airborne Radar or ground-penetrating Radar measurements or FLI-MAP reflectivity maps. These could also be used for filtering purposes or possibly for the detection of buried channels.

More research could be done into the exact position and size of sand layers that are detected using drillings from the DINOLoket. This could result in a number of map products: the depth of the sand layer for each detection area with a probability map indicating what the chance is that this sand layer is at that depth.

A common difference between the classification results and the palaeogeographic map is the channel width. Research could be done into the causes of these differences.

# Bibliography

- [1] J.R.L. Allen. A review of the origin and characteristics of recent alluvial sediments. *Sedimentology*, (5):89–101, 1965.
- [2] N.J. Bakker. TOP10NL in ontwikkeling. pages 48–54, 2002.
- [3] H.J.A. Berendsen. *Landschap in delen - Overzicht van de geofactoren. Fysische geografie van Nederland*. Koninklijke Van Gorcum, 2005.
- [4] H.J.A. Berendsen and P.C. Beukenkamp. Het landschap ten zuiden van Utrecht. Een fysisch-geografische excursie. *De Nieuwe Geografenkrant*, 7(3):23–26, 1983.
- [5] H.J.A. Berendsen and E. Stouthamer. *Palaeogeographic development of the Rhine-Meuse delta, The Netherlands*. Koninklijke Van Gorcum, 2001.
- [6] H.J.A. Berendsen and K.P. Volleberg. New prospects in geomorphological and geological mapping of the Rhine-Meuse Delta – Application of detailed digital elevation maps based on laser altimetry. *Netherlands Journal of Geosciences*, 86(1):15–22, 2007.
- [7] P.J. Besl and R. Jane. Invariant surface characteristics for 3D object recognition in range images. *Computer Vision, Graphics Image*, (33):33–80, 1986.
- [8] L. Bian. *Multiscale Nature of Spatial Data in Scaling Up Environmental Models*. CRC Press, Inc., 1997.
- [9] Website Dorling Kindersley Books. <http://www.dorlingkindersley-uk.co.uk>. Visited 31 January 2009.
- [10] P.A. Burrough and R. McDonnell. *Principles of Geographical Information System*. Oxford, Oxford University Press, 1998.
- [11] P.A.B. Burrough. *Principles of Geographical information Systems for Land Resources Assessment*. Oxford, Clarendon Press, 1986.
- [12] J.B. Campbell. *Introduction to Remote Sensing*. The Guilford Press, 2007.
- [13] J. Cohen. A coefficient of agreement for nominal scales. *Educational and Psychological Measurement*, (20):37–46, 1960.
- [14] B.A. den Toom, N. van der Ven, and A.J. Osté. Toelichtingen bij de peilbesluiten in de Alblasserwaard. Technical Report 13010213, Ingenieursbureau BCC bv, 2005.

- [15] E.R. Dougherty and R.A. Lotufo. *Hands-on morphological image processing*. The International Society for Optical Engineering, 2003.
- [16] R. Dubayah, E.F. Wood, and D. Lavallée. *Multiscaling Analysis in Distributed Modelling and Remote Sensing: An Application Using Soil Moisture. Scale in Remote Sensing and GIS*. Lewis Publishers, 1997.
- [17] Persbericht Provincie Flevoland. <http://www.nieuwsbank.nl/inp/1999/09/0923F141.htm>. Visited 27 October 2007.
- [18] Website FLI-MAP. <http://www.flimap.com>. Visited 30 October 2008.
- [19] Website GBKN. <http://www.gbkn.com>. Visited 10 March 2008.
- [20] R. Keshet. Dilation.png, July 2008. Created in Inkscape.
- [21] R. Kranenburg. Is er toekomst voor Flevoland? *Geografie*, (3):6–13, 2006.
- [22] J.R. Landis and G.G. Koch. The measurement of observer agreement for categorical data. *Biometrics*, (33(1)):159–74, 1977.
- [23] H. Leenaers. *Book: Onderste Boven*. TNO, 2005.
- [24] L.B. Leopold and M.G. Wolman. River Channel Patterns: Braided, Meandering and Straight. *U.S. Geological Survey Prof. Paper*, (282-B):39–45, 1957.
- [25] W.G. Madow and L.H. Madow. On the Theory of Systematic Sampling. *I. Ann. Math. Statist.*, 15(1):1–24, 1944.
- [26] M. Molenaar. *An introduction to the theory of spatial object modelling for GIS*. Taylor and Francis Inc., 1998.
- [27] J.R. Mulder. Op zoek naar resten van oude geulen in de kronkelwaard van de Hondsbroekse Pleij bij Westervoort, een veldbodemkundig en historisch-geografisch onderzoek. Technical Report Alterra-rapport 1170, Alterra, Wageningen UR, 2005.
- [28] N.A. Nabih. Algebraic error analysis for surface curvature and segmentation of 3-D range images. *Pattern Recognition*, 23(8):807–817, 1990.
- [29] G. Oude Essink and R. Stuurman. Onderzoek naar de effecten van de Hoofdwegverbinding A6-A9 op zoet-zout grondwater nabij het Naardermeer. Technical Report 2006-U-R0072/A, TNO Bouw en Ondergrond, Utrecht, 2006. <http://www.tno.nl/downloads/1.60%20Effecten%20hoofdwegverbinding%20A6-A9.pdf>.
- [30] H.E. Reineck and I.B. Singh. Genesis of laminated sand and graded rhythmites in storm-sand layers of shelf mud. *Sedimentology*, 18(1-2):123–128, 1971.
- [31] J. Serra. *Image Analysis and Mathematical Morphology*. Academic Press, London, 1982.
- [32] D.G. Smith. Anastomosed fluvial deposits: modern examples from Western Canada. *Special Publication of the International Association of Sedimentologists*, (6):155–168, 1983.
- [33] Website Safe Software. <http://www.safe.com>. Visited 21 March 2008.
- [34] M. Spivak. *A comprehensive introduction to differential geometry (3rd edition)*, volume 3-4. Publish or Perish Press, 1999.

- [35] E. Stouthamer. *Sedimentary products of avulsions in the Holocene Rhine-Meuse delta, The Netherlands*. *Sedimentary Geology* 145, 2001.
- [36] P.H. Swain and S.M. Davis. *Remote Sensing: The Quantitative Approach*. New York: McGraw Hill, 1978.
- [37] P.J.G. Teunissen. *Adjustment theory*. Delft University Press, 1999.
- [38] Website TOP10NL. <http://www.kadaster.nl/top10nl>. Visited 29 January 2009.
- [39] T.E. Törnqvist. Fluvial sedimentary geology and chronology of the Holocene Rhine-Meuse delta, the Netherlands. *National Geographical studies*, (166):169, 1993.
- [40] T. van der Ven. *Poster presented at NAC 9*. 2008.
- [41] R.M. van Heerd, E.A.C. Kuijlaars, M.P. Teeuw, and R.J. van t Zand. Productspecificatie AHN 2000. Technical Report MDTGM 2000.13, Rijkswaterstaat, Data-ICT-Dienst, 2000. <http://www.ahn.nl>.
- [42] F. van Hemmen, D. Bekius, and E. Heunks. Schone slaper - Hollands hoop in bange dagen. Cultuurhistorisch advies voor de verbetering van de Diefdijklinie. Technical Report 1531, Raap, Amsterdam, 2007.
- [43] Website Vijverhof. <http://www.de-vijverhof.nl/index.php?p=historie>. Visited 29 January 2009.
- [44] A. Weiss. *Topographic Position and Landforms Analysis*. 2001. Poster presented at the ESRI User Conference, San Diego, July 9<sup>th</sup>.



# Appendix A - Workflow Diagram

The workflow diagram can be seen in Figure 7.1. It illustrates the different processing steps and the order in which they are applied.

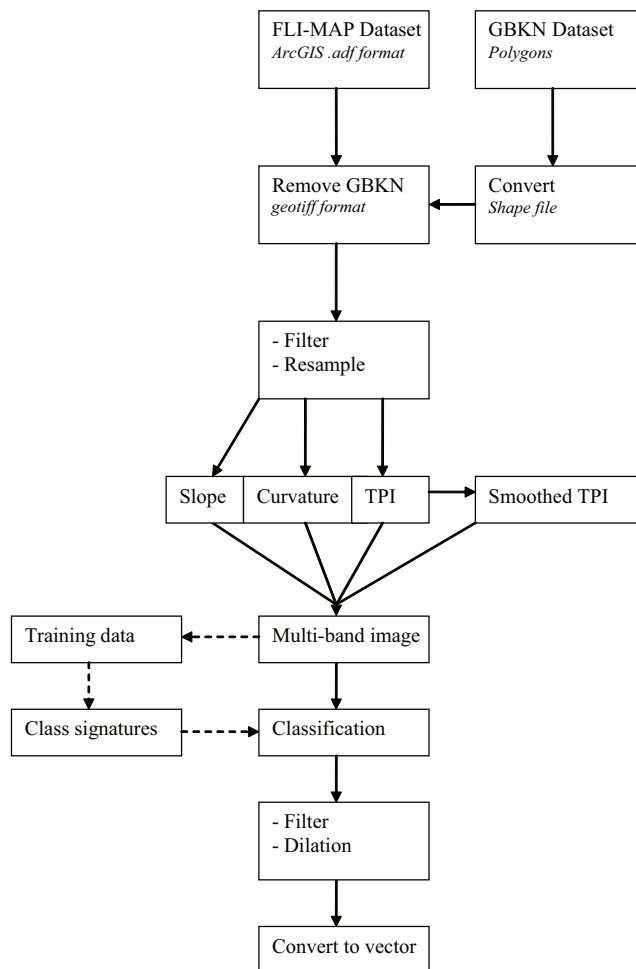


Figure 7.1: Workflow diagram.



# Appendix B - Software

Various programs and programming languages were used during the work on this thesis, depending on the capabilities or requirements. Each of these are explained in this appendix in the order they were used.

- FME (Safe Software)  
Used for adapting the GBKN data. This program is used because it can cope with large amounts of spatial data and has a wide variety of integrated processing features.
- ArcGIS (ESRI)  
Primarily ArcGIS was used as interface for programming algorithms. It was also used for analysis, various calculations and to visualise the datasets.
- Python 2.5.2 (Python Software Foundation, open source license)  
Used for programming the algorithms and is implemented in a toolbox in ArcGIS. This programming language was used because it is supported by ArcGIS, but mainly because it is open source software. Due to this, every aspect can be adapted and there are many examples and help sources available online.
- Additional packages for Python:
  - *GDAL 1.5.2*  
Geospatial Data Abstraction Library (GDAL) is used to let Python work with geographical data. It is a translator library for raster geospatial data formats and it comes with a variety of useful command-line utilities for data translation and processing.
  - *Numpy 1.1.0*  
NumPy is the fundamental package needed for scientific processing using Python. It has linear algebra capabilities and is needed to work with matrices.
  - *PyWin32 2.1.1*  
This package contains the Python for Windows Extensions, used to provides access too many of the Windows API's. An 'Application Programming Interface' (API) is an interface in computer science that defines the ways by which an application program may request services from libraries and/or operating systems.
- Matlab (The MathWorks)  
Initially Matlab was used to test different detection algorithms. After testing these algorithms were implemented using Python. Matlab was also used for importing and the analysis of the DINOLoket drilling data.



- Flex Viewer (ESRI)  
This online GIS viewer makes it possible to visualize geographical data on every computer with an internet connection and Adobe Flash player. The data are accessible through internet, processing is performed on the server using a web map service and the results are converted to a picture and sent to the viewer of the user. The online Flex viewer used in this research is based on the free Adobe Flex framework.
- L<sup>A</sup>T<sub>E</sub>X  
This typesetting program was used for writing the report.

# Appendix C - Borehole description example

**Table 7.1:** *Borehole description example part 1*

ALGEMENE GEGEVENS BORING	
NITG-nummer:	B38E0309
X-coördinaat (m):	120837
Y-coördinaat (m):	437648
Coördinatensysteem:	RD2000
Plaatsnaam:	Onbekend
Provincie:	Zuid-Holland
Kaartblad:	38E
Bepaling locatie:	Onbekend
Maaiveldhoogte (meter t.o.v. NAP):	-0.95
Bepaling maaiveldhoogte:	Onbekend
Boormethode:	Onbekend
Einddiepte (meter beneden maaiveld):	10.00
Datum boring:	null
Eigenaar:	Onbekend
Uitvoerder:	Onbekend



# Appendix D - Processing time

Table 7.3 shows the approximate amount of processing time necessary to perform significant operations on the data. The term “processing time” refers to the time the computer is calculating the result and does not include any programming, visualisation or other operation on the data. Processing was performed on a desktop computer with an Intel® Core™2 Duo 2.13 GHz processor and 1 GB of RAM, running on Windows XP.

**Table 7.3:** *Processing time*

<b>Filtering:</b>	
GBKN	30h
Variance	1h
Density	1h
<b>Detection:</b>	
Slope	6h
Curvature	6h
TPI	18h
Smoothing	1h
Forming composite-band	1h
Classification	$\frac{1}{2}$ h
Result cleanup	1h
<b>Validation:</b>	
Import DINO-data	$\frac{1}{4}$ h



# Appendix E - Discarded Processing techniques

Besides the main detection algorithm explained in this thesis, numerous other processing/detection techniques were tried. The reasons why they could not be used are explained below.

- **Hydrological model/flow direction**

Hydrological models are used to simulate the flow of river or stream and to calculate water accumulation, drainage patterns and directions. There are a number of steps that need to be performed to apply this method.

- The first step is to make the dataset usable for hydrodynamic calculations. This is done by inverting the heights so that the buried channel deposits are lower than the surrounding landscape. This will let the simulated water accumulate towards buried channel deposits instead of running off.
- Fill in holes/sinks in the dataset (because of earlier inversion, they are in fact bumps in the real landscape). This is necessary because sinks do not drain and lead to accumulation at a single point. In the real world this would form a pond or a lake that eventually spills over again. But most hydrological models have difficulty with this process, so the solution is to fill sinks before running simulations.
- Calculate flow direction. This process creates a raster of flow direction from each cell to its steepest downslope neighbour.
- Flow accumulation is calculated from flow directions. It calculates the number of grid cells that drain into an individual grid cell.
- Finally, the drainage lines can be calculated and the stream can be divided into segments.

Hydrological models can not be used for the detection of buried channel deposits because there is not enough elevation difference in the dataset. The surface of the Alblasserwaard is very flat compared to the mountainous terrain where this method is usually applied. Also the surface features of buried channel deposits are often disturbed, this prohibits the calculation of the flow direction (step 3) and the corresponding accumulation network (step 4).

- **Fourier Image Analysis**

Fourier Image Analysis was already applied as a filtering technique. Height data were converted into the frequency domain using a 2D Fourier transformation. The assumption is that buried

channel deposits cause low frequencies and the noise is contained in the high frequency part. So in the frequency domain the spectrum of the higher frequencies was removed, the middle and low frequencies were preserved. An inverse 2D Fourier transformation was then performed to transform the data back into the spatial domain (the image). Looking at the result, this method did not only filter but also introduced a lot of new errors. Further analysis of the lower frequency data showed that the signals of buried channel deposits are contained in a broad spectrum and could not be singled out.

- **Directional Cross-sections + Gaussian fitting**

This method was developed by looking at height profiles in different direction from a single point. Possibly a classification could be made by comparing these profiles with characteristics of profiles from buried channel deposits. Profiles in 4 directions were fitted with Gaussian curves (modeled after a Gaussian distribution). The parameters of these curves were used for classifying each point. There are a number of reasons why this method did not work and is probably not the best suited method for detection buried channel deposits. First of all, determining the values for classification based on the fitted parameters was very difficult. Secondly, local disturbances in the elevation data caused many errors in parameters estimation, leading to incorrect classification. And lastly, this method is highly computational intensive. It could have taken weeks or months to process all the data.

## **Appendix F - Elevation dataset of the Alblasserwaard**



## Appendix F - Elevation dataset of the Alblasserwaard

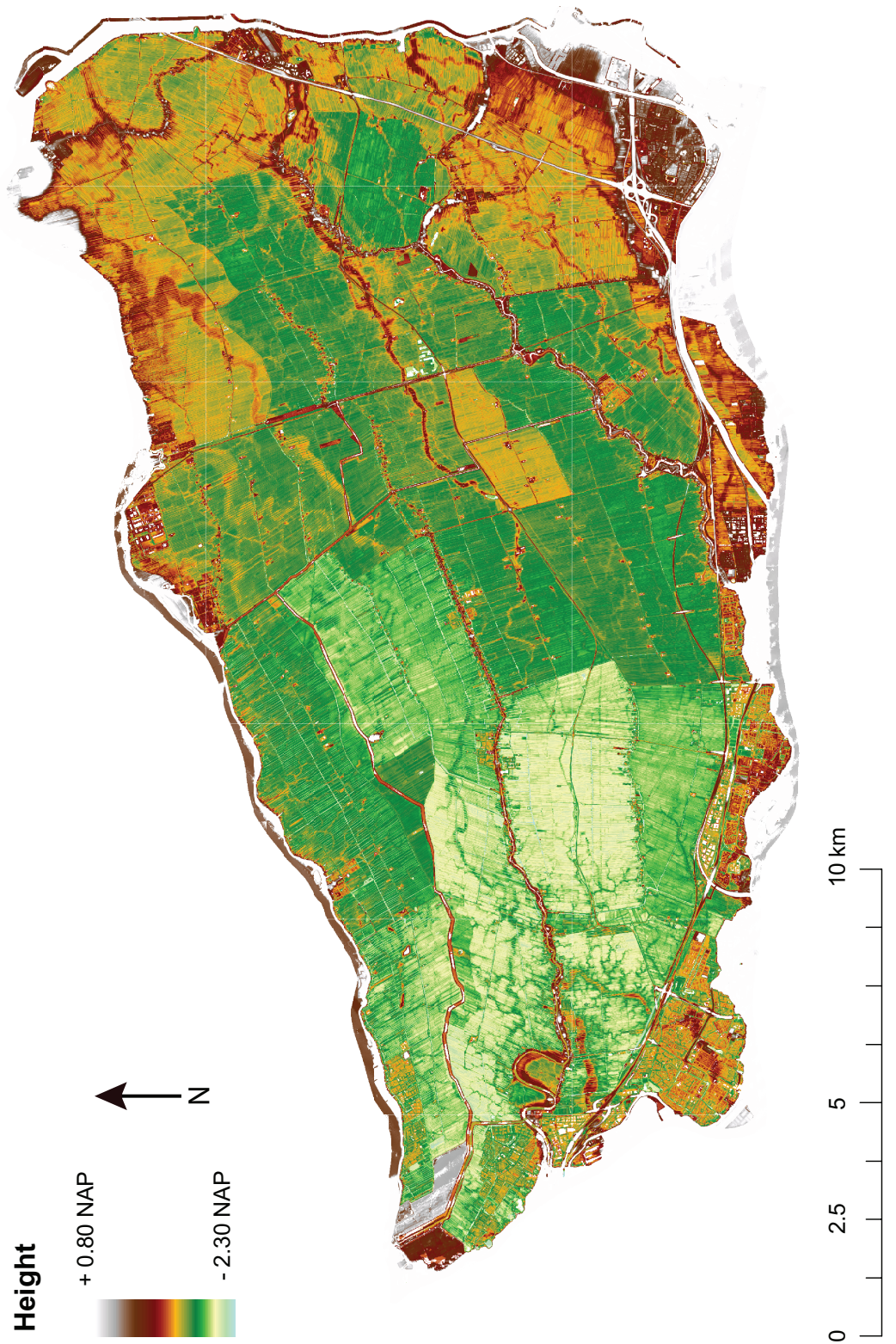


Figure 7.2: Elevation dataset of the Alblasserwaard.

# Appendix G - Field measurements

One day of fieldwork was performed in the Alblasserwaard. In total 3 shallow drillings were performed using a gouge (in Dutch: "guts"). In Figure 7.3 and Figure 7.4 some photo's can be seen.



**Figure 7.3:** *The location in the Alblasserwaard where the fieldwork was performed.*



**Figure 7.4:** *Close-up of a gouge used for the drillings. You can see a cross-section of approximately 0.5 m of soil (in this case clay).*

Point 1:  
N 51° 51.271'  
E 004° 41.920'

Point 2:  
N 51° 51.199'  
E 004° 41.930'

Point 3:  
N 51° 51.160'  
E 004° 41.862'