1 Preface

This document forms a thesis proposal submitted to the Delft University of Technology in partial fulfillment of the requirements for the degree of Master of Science in Geomatics. This proposal will introduce the reader to the defined topic and research objectives in Section 2, will list and describe the relevant related works in Section 3, presents the materials and methods in Section 4, and will finally give a time planning of the entire project in Section 5.

2 Introduction

Drainage networks affect hydrology in many types of environments; floodplains (Cazorzi et al., 2013), peatlands (Holden et al., 2004), coastal wetland zones (Poulter et al., 2008), and mountain regions (Cavalli et al., 2013). In agrarian areas, artificial drainage networks have been reported to support ecological hot-spots and function as ecological corridors (Carpentier et al., 2003; Croxton et al., 2005; Pita et al., 2006). Furthermore, the drainage network greatly impacts the environmental functioning in these areas through the constraining of water flow paths in the landscape (Bailly et al., 2011), controls water and pollutant fluxes (Bouldin et al., 2004; Dages et al., 2008; Cazorzi et al., 2013), and influences groundwater hydrology (Dages et al., 2008) and hydrological response during flood events (Cazorzi et al., 2013). Agricultural ditch drainage networks have been reported to affect the hydrological environment in four different ways: (1) altering of surface runoff (Carluer and Marsily, 2004; Dunn and Mackay, 1996; Levavasseur et al., 2012), (2) drainage of groundwater (Dunn and Mackay, 1996), (3) recharge of groundwater (Carluer and Marsily, 2004), and (4) transport of water downstream (Carluer and Marsily, 2004; Dunn and Mackay, 1996; Levavasseur et al., 2012).

Artificial drainage networks (Figure 1) in agrarian landscapes consist of connected linear features such as channels, culverts, and reshaped gullies (Bailly et al., 2011). Together these linear features form a network of structures which transit water from the fields into larger canals (Bouldin et al., 2004). An up-to-date and detailed recognition of the channel network is crucial for landscape management issues such as water resources management (Cavalli et al., 2013). The storage capacity within the network plays an important role in designing drainage channels and pumping stations (Malano and Hofwegen, 1999), and the assessment of network storage capacity is a crucial tool in flood management as it can identify areas which are potentially vulnerable to floods (Cazorzi et al., 2013). A correct network characterization containing large-scale and up-to-date positioning and geometry of the drainage network is beneficial for the programming of measures, which can guarantee safety from flooding (Cazorzi et al., 2013).
Figure 1: (a) An examplary cross section of a drainage channel. The important characteristics are the ‘center line of the channel’, ‘the water/land boundary’, the ‘bankfull boundary’, and the ‘storage capacity’. Banks typically have a slope between 1/2 - 1/10, or 1/0, and have equal slopes only in idealized situations. (b) Together, these drainage channels form a network that drains the fields and directs the water into larger channels or rivers.
Manual field observations and photo-interpretations are traditional methods for the delineation of the drainage network. Manual field observations are limited by human resources and money constraints (Gandolfi and Bischetti, 1997), and involve the subjective judgement of the observer (Gandolfi and Bischetti, 1997). Similarly, network detection from aerial images is subject to the same subjectivity. Furthermore, this technique has problems with obscuration and misleading effect of the canopy, the image scale, and the presence of distortions and shadows. Good quality large scale images, and a large amount of work, are needed to correctly delineate the drainage network using aerial images (Gandolfi and Bischetti, 1997).

Light Detection And Ranging (LiDAR) has become an accepted means of acquiring topographic data because of short data acquisition and processing times, relatively high accuracy and point density, and reductions in acquisition costs (Flood and Gutelius, 1997; Hill et al., 2000; Charaniya et al., 2004). Lidar can provide very precise horizontal and vertical information, with the Actueel Hoogtebestand Nederland 2 (AHN2) providing a systematic and stochastic error of maximum 5 cm (van der Zon, 2013). Artificial network dimensions are often of larger width or height than the resolution of topographic data provided by LiDAR sources, thus this kind of data has high potential for the use in mapping of linear landscape elements such as artificial drainage networks (Bailly et al., 2008).

The Netherlands is characterized by mostly low lying lands, of which approximately 55% is sensitive to flooding (ref). To protect against floods, and to maintain an optimal water balance, a good drainage system is thus of the utmost importance. In the Netherlands there are 24 water boards, ‘waterschappen’ in Dutch, which each are responsible of the water resources management in their designated areas. One of these water boards, ‘Hoogheemraadschap De Stichtse Rijnlanden (HDSR)’, uses the SOBEK software suite by Deltares for hydrologic modelling in their working area. A correct and up-to-date characterization of the channel network in terms of reaches, and the cross-sectional dimensions of these reaches, is beneficial for the correctness of the model output. Current methods employed by the HDSR to characterize the channel network are based on stereo-imaging, which is slow, cumbersome, and subjective. The cross sectional dimensions of the individual channels are obtained by manual field observation. The HDSR wishes to have a highly automated method to characterize the channel network and dimensions, which would allow more frequent updates in the future. A LiDAR dataset, the AHN3, is available for the Netherlands which can be used to develop the new technique.

### 2.1 Challenges

The development of the proposed automated method will face a number of challenges. First of all, the exact position of the channel center line depends on the water level at the time of measurement (Figure 2), which can lead to positional inconsistencies between datasets acquired at different moments in time. It is intuitive to use the center line of the water surface, however taking the center line in between the bankfull boundaries would mitigate these possible positional inconsistencies, since the bankfull boundaries are less likely to change between acquisition times. However, it is dependent on the application which definition of the center line is preferable, e.g. for visualisation purposes it may be undesirable to have a center line based on bankfull width (Figure 3). Thus it would be preferable if the automated method is able to identify both. Both the width of the water surface and the width of the bankfull channel have to be obtained to determine these center lines.

Secondly, obtaining the cross sectional dimensions of individual channels (for estimating the storage capacity) minimally requires, next to the width of the water surface and the width of the bankfull channel, obtaining the height difference between the water surface and water surface at bankfull discharge. This would give a rough approximation of the channel cross section,
Figure 2: Schematic cross section of the water surface during high and low water, with channel center indicated in red. The channel center is at a different positions depending on the height of the water surface.

Figure 3: Schematic cross section of the water surface during high and low water, with channel center indicated in red. The channel center at high water (bankfull width) is positioned above what might commonly be referred to as the land surface. For visual display on maps, this might not be desirable.
The characterization of the drainage network is important for the use in water resources management and hydrologic modelling, which is especially true for the vulnerable situation of the Netherlands.

This thesis will identify the possibilities to characterize the drainage network from a LiDAR point cloud dataset, thereby striving for a high level of automation. The result will be a workflow that takes a LiDAR point cloud as input, and accordingly identifies and returns the geographical position and geometry of every drainage channel. The geographical position will be characterized by both the center line of the water surface and the center line of the bankfull surface. The geometry of the drainage channel is formed by finding (along the entire length of the drainage channel) the width of the water surface, the width of the bankfull surface, and the channel banks. The workflow will thus result in the creation of three datasets: (1) 2D center lines of the water surface, (2) 2D center lines of the bankfull surface, and (3) a 3D geometry of the entire drainage network. The three datasets will be topologically sound, i.e. the individual drainage channels will together form a connected drainage network, thus channel topology will be repaired where possible. The resulting datasets will also be checked, and if possible corrected, for positional accuracy. The main research question of this thesis will be:

“To what extent can the position and 3D geometry of the drainage network be automatically identified from a LiDAR point cloud?”

To answer this main research question, the following sub-questions will be answered:

- What are the typical properties of water courses that make them different from the rest of the landscape, and how well can these be extracted from the LiDAR point cloud?
- How can the center lines of the water surface and bankfull surface be extracted from these properties?
• How can the channel banks be extracted from the LiDAR point cloud?

• How can the 3D geometry of the channels be constructed using the water surface, bank-full surface, and channel banks?

• How can topological and positional defects in the 2D and 3D channel network be identified and repaired?

2.3 Scope

This thesis will not deal with pre-processing, filtering, and classification of the LiDAR dataset. The AHN3 LiDAR point cloud will be used as distributed, along with the built-in classification of six classes (ground surface, water, buildings, artificial objects, vegetation, and unclassified). The workflow will be applied to the working area of the HDSR in the province of Utrecht, the Netherlands. A tiled LiDAR dataset will be used, and the method will be tested along the seams of these tiles. Due to the use of terrestrial LiDAR in AHN3, only the 3D geometry of the channel network above the water surface can be obtained.

3 Related Works

Manual observations and photo-interpretations are traditional methods for the delineation of the drainage network. However, these approaches both require a significant amount of human labour and involve the subjective judgement of the operator (Gandolfi and Bischetti, 1997). Furthermore, the use of aerial images for network detection is limited by the obscuration and misleading effect of the canopy, image scale, and presence of distortions and shadows (Gandolfi and Bischetti, 1997). Light Detection And Ranging (LiDAR) has become an accepted means of acquiring terrain elevation data because of short data acquisition and processing times, relatively high accuracy and point density, and reductions in acquisition costs (Flood and Gutelius, 1997; Hill et al., 2000; Charaniya et al., 2004). The use of LiDAR mapping to extract channel-like features has been applied to many disciplines: for the detection of tidal creeks (Lohani and Mason, 2001; Mason et al., 2006), channel heads (Clubb et al., 2014), moraines (Rutzinger et al., 2006b), gullies (James et al., 2007; Eustace et al., 2009; Baruch and Filin, 2011), buildings (Rutzinger et al., 2006a), streams (Cho et al., 2011), coastal structural lines (Brzank et al., 2008), and surface fissures on landslides (Stumpf et al., 2013). The vast majority of the methods developed in these interdisciplinary studies have not been applied to the mapping of (artificial) drainage networks in agricultural areas, however they may still be relevant to the cause of this thesis. Thus, this literature study will not be limited to the extraction of artificial drainage networks, rather it will also consider methods designed for mapping of other channel-like features.

3.1 Mapping of channel-like features

In general, the methods for mapping of channel-like features can be divided into the following five categories: Flow-routing models, thresholding models, High-Level Image Processing models, point-based methods, and other methods.

3.1.1 Flow-routing models

Many methods have been developed which extract terrestrial fluvial channels from Digital Elevation Models (DEMs) using flow accumulation algorithms. Jenson and Domingue (1988) use the method of steepest gradient, whereby the channel network is extracted by finding flow
accumulation values for each cell, and accordingly selecting the pixels for which flow accumulation values exceed a certain selected threshold value. Costa-Cabral and Burges (1994) use a multiple-direction method which allows distributing flow from a pixel among all of its lower-elevation neighbor pixels, which improves on the steepest gradient approaches. Lohani and Mason (2001) applied the flow-routing model by Jenson and Domingue (1988) to LiDAR data of an intertidal zone and found the performance of these methods to be moderate, which they presume can be accounted to the fact that these methods were developed for terrestrial fluvial channels, which posses different morphologic and hydrodynamic characteristics from those of tidal channels. The local aspect, curvature, slope, and drainage density of an intertidal zone generally differ from those of a terrestrial fluvial basin (Lohani and Mason, 2001). Tidal channels are incisions on flat or gently sloping ground (Lohani and Mason, 2001), which gives reason to believe that flow-routing models will also not work well for extraction of artificial drainage networks, since they are morphologically more similar to tidal creeks than to terrestrial fluvial channels. Flow-routing methods assume that slopes along the channel path remain positive, and consider flow-path as beginning near ridges and in relatively high curvature values (Baruch and Filin, 2011). Such assumptions do not apply for artificial drainage networks in flat landscapes such as the Netherlands, and Bailly et al. (2008) notes that studies on artificial drainage network detection in agrarian landscapes showed that drainage algorithms on DTM are not suitable in the case of anthropogenic networks. Furthermore, the methods by Jenson and Domingue (1988) and Costa-Cabral and Burges (1994) were developed for and tested with low-resolution DEMs (+-30m?), while DEMs of LiDAR data can be of much higher spatial resolution (0.5 for AHN3).

3.1.2 Thresholding methods

Chorowicz et al. (1992) compare the elevation of each pixel in a DEM of a mountainous terrain with that of its neighbor. Values are then assigned to the boundaries between these pixels to indicate whether a pixel is higher or lower than the previous point by an amount greater than a certain threshold. Fagherazzi et al. (1999) extracted tidal creeks from a DEM based on elevation and curvature thresholds. Meisels et al. (1995) used multilevel skeletonization to identify pixels having curvature higher than a threshold. Brzank et al. (2005) attempts to segment water from non water areas in coastal regions. They select local height minima below an empiric determined threshold to be used for region growing. They accordingly check whether the mean intensity value of the region is lower than an empiric given intensity value, to find whether the area is water or not. Rutzinger et al. (2006b) used an object-driven analysis with predefined windows and thresholds to extract moraines. Cho and Slatton (2007) use the inverse tangent of a convolution operation on a predefined window to emphasize variations in elevation between stream channels and the surrounding stream banks under a dense forest canopy. This method was however tested on a coarse 1x1m DEM, and is specifically useful for application in deep and wide profile features (Baruch and Filin, 2011). Finally, Cavalli et al. (2013) used a curvature-based and a modified slope-dependent threshold area approach for channel network extraction on the basis of a 4m resolution LiDAR DTM.

The thresholding methods above use pre-defined threshold values and window sizes that reduce their adaptability for detecting features which display significant form diversity (Baruch and Filin, 2011), also this may limit these techniques to use with low-resolution data only (Liu et al. 2015). To counter this limitation, Baruch and Filin (2011) use a multi-scale parameter analysis to detect gullies based on surface curvature parameters, but their method is designed for use in strongly textured terrains. Lohani and Mason (2001) used an adaptive height threshold to locate tidal channels, but their approach is limited to low-resolution data where channel width is on the single-pixel level (Baruch and Filin, 2011).
3.1.3 High-level image processing (HLIP) methods

HLIP methods use the geometry and geomorphology of the landscapes’ linear features to find the channel network (Liu et al., 2015). Mason et al. (2006) presented a semi-automatic multi-level knowledge-based approach. High-gradient pixels were found using edge detectors, and this was followed by edge association, centerline generation, network repair, and channel expansion. Mason et al. (2006) extend upon the method by Lohani and Mason (2001) to also handle high-resolution LiDAR data. Brzank et al. (2008) extracted structural lines in the Wadden Sea by fitting a two-dimensional (2D) hyperbolic tangent curve. Cho et al. (2011) used a composition of geodesic top-hat and bot-hat operations of different sizes to build a morphological profile, thereby recording the image structural information. They comment that it requires significant training and computation. Liu et al. (2015) developed the automated method for extracting tidal creeks (AMETC), which uses a multi-window median neighborhood analysis to enhance depressions in mudflat and marsh environments, a multi-scale and multi-directional Gaussian matched filtering method to enhance width-variant tidal creeks, and a two-stage adaptive thresholding algorithm to segment low-contrast tidal creeks. The method successfully extracted both small and large tidal creeks from the study area. The method is claimed to be weakly dependent on scale, robust, and automatic. Although the HLIP methods give promising results, none of these have yet been applied to artificial drainage networks in agricultural areas.

3.1.4 Point-based classification

Brzank et al. (2008) extract water from non water points in a LiDAR data set by means of a supervised fuzzy logic concept. They group points into single scan lines, after which a membership value of class water is calculated for each point of every scan line based on height, intensity, 2D point density, and angle of incidence of every point. Brodau and Lague (2012) use a multi-scale measure of the point cloud dimensionality around each point. This dimensionality characterizes the local 3D organization of the point cloud within spheres centered on the measured points and varies from being 1D, 2D, to a full 3D volume. By varying the diameter of the sphere, it can be observed how the local cloud geometry behaves across scales. The technique was applied to separate riparian vegetation from ground, and classifying a mountain stream as vegetation, rock, gravel, or water surface. They claim their technique to being fast and accurate, and can be fully automated. LiDAR however has the property of being absorbed by water bodies, thus leading to many gaps in the data, and effectively making use of the previous techniques impossible. Höfle et al. (2009) try to mitigate this problem by modeling the locations of laser shot dropouts based on timestamps of the recorded laser measurements. Accordingly, they apply a seeded region growing segmentation algorithm to the point cloud and modeled dropouts to detect potential water regions. Finally, object-based classification of the resulting segments determines the separation between water and non-water points. Although this is a promising approach, the AHN3 dataset is not shipped with timestamps of the recorded laser measurements. Furthermore, the approach requires significant pre-processing. Toscano et al. (2014) propose using LiDAR intensity data as well as an advanced histogram analysis of LiDAR elevation data to automatically detect and delineate water bodies with very little pre-processing. Since the surface of a water body has the same elevation, a peak at the corresponding elevation can be observed in the histogram. To detect smaller water bodies, LiDAR intensity data is essential.
3.1.5 Other methods

Cho et al. (2011) detected stream channels using mathematical morphology operators in a heavily forested and rugged landscape. Similarly to the thresholding methods, Liu et al. (2015) notes that mathematical morphology methods use pre-defined threshold/mask/structure-elements that reduce their adaptability for detecting features which display significant form diversity, also this may limit these techniques to use with low-resolution data only. Zhu and Toutin (2013) use a two-hierarchy decision-tree classification to discern between six land cover class, among which water bodies. The study however does not specifically aim at channels, but mainly at larger water bodies.

3.2 Mapping of artificial drainage networks

Very few authors have tried to map artificial drainage networks in agricultural areas using LiDAR data. Bailly et al. (2008) proposed a method to find artificial drainage networks in agricultural landscapes. Their methodology uses LiDAR data in three steps. First they estimate elevation profiles from LiDAR points on a set of pre-located sites. Then they perform profile shape description with wavelet transform or a watershed algorithm. Finally, they classify the profiles using classification trees with predictors coming from shape analysis. They claim their method can work well for the automatic detection or characterization of certain agrarian linear features, but not for agricultural ditch networks. Moreover, Cazorzi et al. (2013) state that their method has two important limitations: (1) ditches can be located exclusively at field boundaries, and (2) a geographic database of plot boundaries must be available. Cazorzi et al. (2013) propose a morphometric methodology based on high resolution DTM (1m). They subtract a DTM from a smoothed elevation model representing an approximation of the large-scale landscape forms, thereby achieving an approximation of the local relief. Only the small-scale topographic features are then preserved, representing a map of residual relief. The peak values of this residual relief are then labelled using a thresholding approach, resulting in a boolean map with network and landscape pixels. Due to the use of thresholding, this method has the same problems related to applicability to different forms as outlined in the thresholding methods section.

Passalacqua et al. (2012) test the ability of GeoNet to extract channel networks in flat and human-impacted landscapes using 3m lidar data for the Minnesota River Basin. GeoNet is a recently developed method based on nonlinear multiscale filtering and geodesic optimization for automatic extraction of geomorphic features from high-resolution topographic data (Passalacqua et al., 2010). Passalacqua et al. (2012) developed a new suite of techniques for GeoNet which are specific for flat landscapes, allowing the identification of both natural and artificial channels that comprise the drainage network, as well as differentiate between the drainage network and engineered structures such as roads and bridges. They propose to use curvature analysis to differentiate between channels and manmade structures that are not part of the river network, such as roads and bridges. They suggest that Laplacian curvature more effectively distinguishes channels in these regions compared with geometric curvature. The entire updated GeoNet package is freely available for download.

However, what all of these methods have in common is that they use derivatives of the original LiDAR datasets. They require the generation of a DEM, which is an inherent problem with such datasets since they contain missing data where the water is located. This is caused due to the absorption of lidar signals are absorbed by water, thus generation of DEMs of these parts is inherently difficult. Hydro-breaklines are required to generate DEMs in such situations, however for that the water surface needs to be known (Toscano et al., 2014).
3.3 Extracting the channel geometry

There are relatively few authors which try to extract the geometry of drainage channel networks. Cazorzi et al. (2013) use a semi-automatic approach to estimate some network summary statistics in an agrarian landscape, among which also channel width. The channel width is found simply by evaluating the number of pixels on a raster map classified as ‘channel’ pixels. This method gives an approximation of the true channel width, of which the precision is dependant on the resolution of the DEM. The method does not give an indication of channel height above the water surface. Höflé et al. (2009) separate water from non-water points and accordingly locates the land-water boundary, from which the channel width could be computed. The usability of this technique depends on the availability of timestamps in the measurement data. The technique is unable to find channel height above the water surface. Passalacqua et al. (2012) extract channel cross sections by simply plotting the topographic data corresponding to a certain transect orthogonal to the channel, after which they identify the river banks and geomorphic bankfull water surface elevation. The method by Passalacqua et al. (2012) is the only method found in literature that extracts both channel cross sections as well as geomorphic bankfull water surface elevation, however the effectiveness of the method completely depends on the ability of generating an accurate DEM above water surfaces. No methods were found that are able to identify the 3D channel geometry above the water surface without depending on the correct determination of a DEM.

3.4 Ensuring topological connectedness of the channel network

To generate a well-connected channel network, the fragments of the channels must be joined (Lohani and Mason, 2001). Lohani and Mason (2001) suggest the use of a weighted distance transform, and subsequently thresholding the distance transform at a specific value. This process fills minor gaps in the channels, but fails to join larger channel gaps (Lohani and Mason, 2001). These larger gaps can be filled by using morphological closing (Lohani and Mason, 2001). Cho and Slatton (2007) use a connectivity number to determine whether disconnected stream centerlines should be joined. The criteria for the connecting process are based on distance and on pixel value differences in the original grayscale DEM image. Baruch and Filin (2011) use a two-step process: first a pruning phase which narrows the potential candidates of linking to the most likely pair-sets, whereafter the optimal path is selected by defining the optimal flow path. Mason et al. (2006) also do something with HLIP. Ahmed et al. (2014) compares and evaluates several map construction algorithms using vehicle tracking data. This is aimed at constructing road maps from GPS data, however the use case is very similar to constructing a river network map from MAT points. They assume that input data is a set of tracks, where each track is a sequence of measurements. These measurements consist of a least lat/lon or (x,y)-coordinates. Several different algorithms are described: Point clustering (including k-means and kernel density estimation), incremental track insertion, and intersection linking.

Point clustering These algorithms assume that the input consists of points which are clustered in various different ways to obtain street segments which finally connect to a street map. Edelkamp and Schrödl (2003) employ the k-means algorithm to cluster the input point set, using distance measures, as a condition to introduce seeds at fixed distances along a path. Their algorithm incorporates new algorithms for road segmentation, map-matching, and lane clustering.
4 Materials and Methods

This Section will introduce the materials and methodology used to answer the research question posed in Section 2.2. An introduction to the study area will be given in Section 4.1 and Section 4.2 will introduce the AHN3 LiDAR dataset. Subsequently, the methodologies will be presented to identify the channel network (Section 4.3), to extract the channel geometry (Section 4.4), and to repair the channel network to obtain a valid topological network (Section 4.5). Lastly, Section 4.6 will list the tools which may be required to fulfill the current research.

4.1 Study area

The working area of HDSR is located in the centre of the Netherlands (Figure 5). Four areas were selected which are together representative for the diversity of environments in the HDSR working area; sandy soils, clay soils, peat soils, and an urban area (Figure 6). The water course and drainage network characteristics may differ between these areas due to their different subsoils.
4.2 AHN3 LiDAR dataset

The ‘Actueel Hoogtebestand Nederland’ (AHN) is a dataset provided by the Dutch Ministry of Transport, Public Works and Water Management, containing detailed and precise elevation measurements covering the entire Netherlands. These elevation measurements were obtained using laser altimetry (LiDAR) from airplane or helicopter. The first version of the AHN (AHN1) was acquired between 1996 and 2003, and was mainly meant for water system management purposes. The second version of the AHN (AHN2) was acquired between 2008 and 2013, and is an improvement relative to AHN1 because of higher point density and precision. AHN2 is characterized by a systematic and stochastic error of maximum 5 cm, with an average point density varying between 6-10 points per m$^2$. The AHN2 point cloud is available in an unfiltered form, and a filtered form only containing the points at field level. The AHN3 is a continuation of the AHN2, retaining the characteristics of the AHN2 with respect to systematic and stochastic error and average point density. During the leafless season of every spring (between 2014-2019), a part of the Netherlands shall be acquired. Apart from the year of measurement, the AHN3 differs from the AHN2 in terms of the added classification of points. The contractors responsible for acquiring the measurements have classified every point in AHN3 into one of the five classes: vegetation, ground surface, buildings, water, and artificial. The AHN3 is provided in the compound CRS EPSG:7415, which uses Amersfoort / RD New (EPSG:28992) for (x,y) coordinates and NAP height (EPSG:5709) for the (z) values. The AHN3 is already available for the selected study areas.

4.3 Identifying the geographical location of the drainage network

Many methods were listed in Section 3 which have been used to identify channel-like features. Of these methods, the methods proposed by Liu et al. (2015), Höfle et al. (2009), and Passalacqua et al. (2012) are the most promising. However, Liu et al. (2015) designed their method for the use in tidal creek environments specifically, and the method has not been applied to agricultural drainage networks. Passalacqua et al. (2012) designed their method...
specifically for an agricultural drainage network, although their study area is somewhat different from the study area used here. Both Passalacqua et al. (2012) and Liu et al. (2015) require the generation of a DEM, which is difficult and prone to error when returns from the water surface are missing. Moreover, the conversion of the LiDAR point cloud into a gridded DEM inevitably leads to a loss in precision, decreasing the potential positional accuracy of locating the channel center lines. For this reason, methods that use the original LiDAR point cloud are to be preferred. Höfler et al. (2009) classify the LiDAR point cloud into water and non-water points, and are thereby able to derive the outlines of the water surface, from which the channel line could be extracted. However, their method requires considerable pre-processing and is unable to detect dry water courses. Furthermore, it detects only the water surface, and not the bankfull surface. There is thus the need for the development of a new method, which uses directly the LiDAR point cloud to identify both center lines for wet as well as dry water courses. I propose to investigate two new methods: the Medial Axis Transform (MAT) (Peters et al., 2015), and disjoint hull (a function available in LAStools). Preliminary experiments have shown that both the MAT and disjoint hull approach have potential in identifying drainage channels. Additionally, to serve as comparison and to identify limitations, the method by Passalacqua et al. (2012) will be applied to the study area here.

4.3.1 Medial Axis Transform (MAT)

The MAT is an alternative skeleton-like shape-descriptor that models an object as a union of balls. The MAT is formally defined as the set of maximal balls that are tangent to the surface of a shape at two or more points. The centers of these maximal balls, also called medial balls, form the medial skeletal structure of the object (Peters et al., 2015). The medial axis thus gives a lower dimensional representation of an object (Ma et al., 2012). Ma et al. (2012) proposed a method to produce approximate medial axis points given a set of surface sample points and their corresponding normal vectors. The method works as follows (Peters et al., 2015) (Figure 7):

- Medial balls are found for each point \( p \) in the point cloud by iteratively shrinking a very large ball that is centered along the point’s normal \( n \). 

- At each of these iterations, a point \( q \) is found that is nearest to the ball’s center. At the next iteration, the ball is accordingly shrunk such that it touches both \( p \) and \( q \), and remains centered along \( n \). 

- The iteration continues until the ball’s interior is empty, and there are no closer points to its center than \( p \) and \( q \). 

This algorithm results in the creation of two products: the interior and exterior MAT. The interior MAT is obtained when the normals point outward (such as in figure 2). The exterior MAT is obtained by flipping the normals, essentially resulting in the complement of the space that is occupied by the interior MAT (Peters et al., 2015). Peters et al. (2015) improved the performance of the shrinking ball algorithm by Ma et al. (2012) for LiDAR point clouds which contain significant noise. The algorithm was extended with heuristics that can prematurely stop the shrinking of the medial balls. These heuristics are based on the progression of the separation angle, which is the angle \( pcq \), where \( c \) is the ball’s center. The shrinking of the medial ball of point \( p \) is stopped when:

- The separation angle of the initial ball is below a threshold \( t_a \) 
- The separation angle of a succeeding ball is below a threshold \( t_b < t_a \)
The medial ball for $p$ is selected as the last ball which does not violate both thresholds. The shrinking ball algorithm is simple, fast, robust to noise, and easy to parallelize (Ma et al., 2012; Peters et al., 2015).

A C++ implementation (masbcpp) of the shrinking ball algorithm to approximate the MAT is available from https://github.com/tudelft3d/masbcpp. This implementation was applied to an AHN3 dataset of one of the study areas, and showed promising results (Figure 8a and Figure 8b). The results in Figure 8a and Figure 8b are promising since the drainage channels are well recognizable when looking at the image. The outer MAT of drainage channels forms a ‘center plane’ (Figure 9), which essentially describes the location of the channel center line depending on the height of the water surface, ranging from the height of the water surface at measurement to the height of the bankfull surface. However, the output of the MAT is not readily usable. In uncomplicated areas, every drainage channel has one medial axis plane around which the outer MAT points are centered (Figure 9). However, for more complicated areas where roads and/or buildings are present near drainage channels, identifying the medial axis planes for the actual drainage channels will be far more complex (Figure 10). Additional procedures will need to be developed to extract the MAT planes corresponding to the drainage channels from these areas.

4.3.2 Disjoint hull

The disjoint hull approach is best explained when looking at a close-up of the original AHN3 dataset (Figure 11a). Looking at this close-up it becomes clear that the water surface of the drainage channels in this dataset can visually be very easily identified, which is due to the fact that all these water surfaces have one thing in common: there are no LiDAR measurements located in the image at these areas. This is due to the fact that water absorbs most of the LiDAR signal, and the measurement equipment only measures a return signal of the water areas which are measured at or around nadir. The AHN3 provides a rough classification, and the few water points that were present in the original AHN3 dataset were filtered out in the close-up of the dataset seen in Figure 11a, leading to artificial drainage channels which are devoid of LiDAR measurements. The disjoint hull approach takes advantage of this by computing 2D boundary polygons for the LiDAR point cloud, and identifying interior holes in the resulting polygons. LAStools’ (http://rapidlasso.com/lastools/) lasboundary tool
Figure 8: Preliminary MAT test using C++ implementation (masbcpp). (a) Original AHN3 dataset. (b) MAT result.

Figure 9: Close-up of MAT result for uncomplicated area
is capable of performing these operations in one simple procedure, and the first results are very promising (Figure 11b). However, the disjoint hull approach still needs to be extended with additional procedures to extract the water surface of the drainage channels from the dataset. The shapefile visible in Figure 11b contains many different polygons, each of which either represent ‘islands’ of land or ‘islands’ of water inside the ‘islands’ of land. The drainage channels in between these islands are still empty spaces. The first step would be to extract these empty spaces from the dataset, and turn them into separate polygons. This could be done by taking the difference between a polygon with the size of the extent of the dataset, and the shapefile of the disjoint hulls. To extract the islands of water from the dataset, they should be checked using ‘iswithin’ operators. If the polygon is within another polygon, then this means it is an island of water inside an island of land, else they are islands of land.

**Deriving center lines of water surfaces**  The above workflow would extract the channels as polygons of different widths, but does not yet create a center line of the water surface. One way to solve this may be by converting the dataset to a grid and subsequently thinning, however this will inevitably decrease the accuracy of the dataset through vector-raster conversions and vice-versa. Another possible procedure is described here: [http://www.ian-ko.com/resources/howto.htm#Create%20street%20center%20lines%20from%20cadastral%20polygons](http://www.ian-ko.com/resources/howto.htm#Create%20street%20center%20lines%20from%20cadastral%20polygons). The procedure would use the original islands of land resulting from the disjoint hull approach, and convert the (merged) polygons to polylines after which Thiessen polygons are created. The Thiessen polygons are then dissolved using the attributes of the original island polygons, which, after converting to polylines, results in the center lines of the drainage channels. The entire procedure is implemented in the ET GeoWizards extension ([http://www.ian-ko.com/ET_GeoWizards/gw_main.htm](http://www.ian-ko.com/ET_GeoWizards/gw_main.htm)) for ArcGIS. Another procedure would be to use the boost geometry package ([http://www.boost.org/doc/libs/1_60_0/libs/geometry/doc/html/index.html](http://www.boost.org/doc/libs/1_60_0/libs/geometry/doc/html/index.html)) to perform a 2D skeletonization.

### 4.3.3 Discussion

The MAT is a promising method which can theoretically be used to extract the center, or ‘medial’, lines of both water and bankfull surface. Moreover, it forms a medial plane which could be used to extract the channel center line for any given water height between the height of the
Figure 11: Disjoint hull approach using LAStools’ lasboundary. (a) Original AHN3 dataset. (b) Disjoint hull results. The dataset is now converted to polygons, and the channel drainage boundaries have been well captured. ‘Islands’ of water have been identified as holes in the polygons.
water surface at the time of measurement, and the height of the bankfull surface. However, subsubsection 4.3.1 described that it is difficult to identify the MAT planes that actually belong to drainage channels. On the other hand, the disjoint hull approach forms a relatively simple procedure to extract water surfaces from the LiDAR point cloud, but is unable to extract the bankfull surface. Furthermore, the disjoint hull approach also cannot be used to locate dry channels. Therefore, this thesis proposes to combine the disjoint hull and MAT approaches. A method will be formed which uses the disjoint hull to extract the water surface, and subsequently identifies all MAT planes which are positioned above this water surface. The center line of the water surface can be derived using the ET GeoWizards extension for ArcGIS or Boost geometry package, and the center line of the bankfull surface can be derived by taking the upper points in the MAT plane, and connecting those in the (x,y) dimension.

This workflow would result in the center lines of the wet channels, but not yet of the dry channels. To get the dry channels, all the previously identified MAT planes for the wet channels will be masked out of the MAT results. This should clean up the dataset, making the dry channels easier to recognize. The MAT planes belonging to the dry channels can then be identified by grouping MAT points which together form a plane of significant length (Figure 12). Points then have to be connected to form a channel center line. Road map construction algorithms such as described in Ahmed et al. (2014) may be used here. The bankfull center line can be identified in the same way as for the wet channels. Since these channels are dry, no water surface center line can be identified.

4.3.4 Proposed workflow

The workflow (using a combination of disjoint hull and MAT) to identify the center lines for wet and dry channels is proposed as follows:

**Wet channels**

1. Use the disjoint hull approach to extract the water surfaces.
2. Derive center line of the water surface using the ET GeoWizards extension for ArcGIS or Boost geometry package.
3. Identify all MAT planes which are positioned above these water surfaces.
4. Derive center line of the bankfull surface by connecting the upper points in the MAT plane in the (x,y) dimension.

**Dry channels**

1. Use the disjoint hull approach to extract the water surfaces.
2. Identify all MAT planes which are positioned above these water surfaces.
3. Mask the MAT planes belonging to the water surfaces from the MAT results.
4. Identify the MAT planes which belong to the dry channels by grouping MAT points which together form a plane of significant length.

4.4 Extracting the 3D geometry of drainage channels

The goal is to extract the 3D geometry of every drainage channel, which requires getting at least the water surface, the bankfull surface, and a characterization of the channel banks. The geometry of the channel below the water surface cannot be obtained, since no information is available due to the absorption of red LiDAR in water.
4.4.1 Obtaining the 3D geometry of wet channels

Extraction of the water surface can be done using the disjoint hull approach as indicated in Section 4.3.2, which leads to polygons of the water surfaces at time of measurement. Then, the following methods can be used to obtain the channel banks and bankfull surface:

**Medial Axis Transform** The outer coordinates of the medial axis transform form a plane of which at least the lower edge is located above the identified water surface. This relation can be used to find for every water surface, the MAT plane which belongs to the corresponding drainage channel. This is useful since the MAT plane is constructed using the points on the banks of the drainage channel, and these points can be stored in the process. Thus, by finding the MAT plane which belongs to a specific water surface, also the LiDAR points which make up the banks of the corresponding drainage channel can be identified. The points on the channel banks combined with the water surface, together can form the 3D geometry of the drainage channel. There is no need to separately extract the bankfull surface, since it is implicitly stored in the geometry of the channel banks.

**Gridded DEM generation for curvature analysis** Another method would be to generate a DEM using the previously identified hydro-breaklines (which is basically what the disjoint hull method does). This DEM can then be used to perform a curvature analysis, and find the location where the curvature exceeds a certain threshold, indicating the border between channel and field. The 3D geometries of the channels can then be extracted from the DEM by a masking approach. This method may work well when channel banks are steep, and thus can be distinctly separated from the surrounding fields. However, this method will work less well when channel banks have little curvature, since it will then be difficult to find this distinction between channel bank and surrounding field. Furthermore, the method is subjective due to the need of setting a threshold curvature value, and the resulting accuracy will depend on the resolution of the DEM.

**Discussion** Out of these two methods, the medial axis transform is to be preferred due to its objectivity and higher potential accuracy. The method does not require the production of an intermediary DEM, and as such may be faster and more efficient. Also the method requires less thresholding. For comparison, both methods will be employed in this thesis.

4.4.2 Obtaining the 3D geometry of dry channels

Obtaining the 3D geometry of dry channels can be done using the same methods as for wet channels. Also here the medial axis transform is the preferred method of choice. The difference is that there is no need to find the outer medial axis planes which belong to the water surfaces, since there is no water surface available. For the dry channels, the corresponding outer medial axis planes have already been identified when deriving the channel center lines. The 3D geometries of wet channels can be formed solely based on the LiDAR points which make up the channel banks, since there is no water surface occluding the lower parts of the channel.

4.5 Repairing the topological network

For hydrological modelling purposes it is important that the drainage network is topologically valid. This means that channels which form a connected network in reality, should also be connected in the derived channel network. Different approaches of network repair will be needed for wet and dry channels, due to their different methods of acquisition.
4.5.1 Repairing topology of wet channels

The entire network of wet channels can, in theory, be extracted using the disjoint hull approach. Thus, the resulting topological network based on wet channels should be complete. However, there are situations where channels are in reality connected underneath bridges or by pipes, which cannot be detected in the LiDAR measurements. These will show up as disconnected channels in the disjoint hull results (Figure 11b). It is unknown whether these channels are really connected in reality, and thus connecting them would be guess work. Channel network repair will be performed on these channels, but the repaired sections will be marked as uncertain. Lohani and Mason (2001) connect channel fragments using a weighted distance transform, thereby thresholding the distance transform to a specific value. They claim this process fills minor gaps in the channel network. Since the gaps in the network in this case will be small, this relatively simple method may suffice and will be used here to repair the channel network.

4.5.2 Repairing the topology of dry channels

Dry channels will be acquired using MAT results, and the resulting dry channel lines will thus differ from the ones obtained for wet channels. Generally, these channels will be more fragmented (Figure 12), which requires network repair. Since gaps will only be minor, the weighted distance transform (Lohani and Mason, 2001) will most likely be sufficient here. This form of network repair will only occur between fragmented dry channels. The network repair will also be performed to connect dry channels to wet channels, however for similar reasons as described above, these repaired channel sections will be classified as uncertain.

4.5.3 Repairing 3D topology

4.6 Tools

The following tools may be required to fulfill the current research:
• masbcpp ([https://github.com/tudelft3d/masbcpp](https://github.com/tudelft3d/masbcpp)) for computing MAT.

• An adaptation of masbcpp ([https://github.com/tudelft3d/masbcpp](https://github.com/tudelft3d/masbcpp)) may be required to compute the reversed MAT.

• Pointio ([https://github.com/Ylannl/pointio](https://github.com/Ylannl/pointio)) for converting from LAS to npy format.

• LAStools ([http://rapidlasso.com/lastools/](http://rapidlasso.com/lastools/)) for working with LAZ and LAS files, e.g. tiling, filtering.

• ArcGIS ([https://www.arcgis.com](https://www.arcgis.com)) for GIS operations.


• QGIS ([http://www.qgis.org](http://www.qgis.org)) for GIS operations.

5 Planning

A GANTT chart was constructed to show the planned phasing of the project and deadlines for presentations and delivery of reports. The chart can be viewed in Figure 13.
Figure 13: GANTT chart indicating the planned project phasing and deadlines.
References


N. van der zon. Kwaliteitsdocument ahn2 (versie 1.3). May 2013. URL [http://www.ahn.nl/binaries/content/assets/hwh---ahn/common/wat+is+het+ahn/kwaliteitsdocument_ahn_versie_1_3.pdf](http://www.ahn.nl/binaries/content/assets/hwh---ahn/common/wat+is+het+ahn/kwaliteitsdocument_ahn_versie_1_3.pdf).