AUTOMATIC DETECTION OF BURIED CHANNEL DEPOSITS FROM DENSE LASER ALTIMETRY DATA

B.M.J. Possel[a], R.C. Lindenbergh[a,*], J.E.A. Storms[b]

[a]Dept. of Remote Sensing, Chair of Optical and Laser Remote Sensing, Delft University of Technology – Boudewijn@Possel.nl, R.C.lindenbergh@tudelft.nl
[b]Dept. of Civil Engineering and Geosciences, Section of Applied Geology, Delft University of Technology – J.E.A.Storms@tudelft.nl
http://www.possel.nl

KEY WORDS: LIDAR, structural classification, shallow subsurface, channel deposits

ABSTRACT:

The formation of the current Rhine-Meuse delta mainly took place during the last 12 000 years. Consecutive avulsions, i.e. sudden changes in the course of river channels, resulted in a complicated pattern of sandy channel deposits, surrounded by peat and clay. Knowledge of this pattern is not only interesting from a geohistorical viewpoint, but is also essential when planning and maintaining constructions like roads and dikes. Traditionally, channel deposits are traced using labor intensive soil drilling. Channel deposits are however also recognizable in the polder landscape by small local elevation changes due to differential compaction. The purpose of this research is to automatically map channel deposits based on a structural analysis of high resolution laser altimetry data. After removing infrastructural elements from the laser data, local feature vectors are built, consisting of the attributes slope, curvature and relative elevation. Using a maximum likelihood classifier, 75 million gridded laser points are divided into two classes: buried channel deposits and other. Results are validated against two data sets, an existing paleographic map and a set of shallow drilling measurements. Validation shows that our method of channel deposit detection is hampered by signal distortion due to human intervention in the traditional polder landscape. Still it is shown that relative young deposits (4 620 to 1 700 years Before Present) can be extracted from the laser altimetry data.

1. INTRODUCTION

During the Holocene (approximately 12 000 years - present), much of the western and central part of the Netherlands was aggrading, as active river systems (Rhine and Meuse) transported sediments from the hinterland to the coastline. In combination with sea level rise this resulted in a Holocene sediment sequence of up to 20 meter thickness. As river channels consist predominantly of sand while the adjacent floodplains were dominated by clay deposition and peat formation, a strong grain-size partitioning occurred. Furthermore, frequent shifts in channel location due to avulsions, resulted in a complex subsurface of clay/peat dominated floodplain deposits laterally and vertically alternating with sand-rich channel areas, (Allen, 1965).

Currently, buried channel deposits are recognizable in the landscape, basically due to a process called topographic inversion. This occurs when floodplain deposits on the sides of the buried channels compact at a higher rate than the channel sand itself. At the surface this results in an area with a higher elevation at the locations of buried channel deposits. Note that the sand remaining from an abandoned channel may not start directly at the surface: channels abandoned relatively long ago may meanwhile have been deeply covered by floodplain deposits. The maximum height differences between the buried channels and the surroundings are in the order of a meter for relative large and young channel deposits. To some extent it holds that the thinner and older, that is, deeper the channel deposits, the smaller also the height difference.

Knowledge of the location of these channel deposits is essential when planning and maintaining large construction works as motorways and dikes, (Munstermann et al., 2008). Abrupt and unidentified changes in the subsurface may lead to unexpected differences in compaction, which may lead again to damaged or uneven road surfaces or even to failing dikes.

Traditionally, mapping of the shallow subsurface of the Rhine-Meuse delta is based on soil drillings. A large effort has been made by Dutch Utrecht University: Based on more than 25 years of field research using over 200 000 manual boreholes a paleogeographic map is composed, (Berendsen and Stouthamer, 2001), see also Fig. 1, right. As the drillings require a large amount of manual labor in the field, not the whole Rhine-Meuse delta has been covered in the same amount of detail. Also necessarily some interpretation and interpolation steps were involved in composing the map, which may have introduced local anomalies.

LIDAR data is being used more and more to reveal and highlight morphological and archaeological structures that are hardly visible. In archaeology, LIDAR data has revealed burial mounts, (Kakiuchi and Chikatsu, 2008), Celtic field systems, (Kooistra and Maas, 2008, Humme et al., 2006), and other earthwork features, (M. Doneus and Jammer, 2008). Spatial scales characteristic for a certain landscape type, like small scale roughness, could be identified by a spectral analysis of airborne laser scanning data, (Perron et al., 2008). Previous studies, (Berendsen and Volleberg, 2007, Munstermann et al., 2008), showed that also buried channel deposits can be visualized using airborne laser data from the AHN (Actueel Hoogtebestand Nederland) archive.

In this research it is considered if it is possible to systematically map channel deposits from second generation, high resolution AHN-2 data. In 2012 for every 50 cm grid point in The Netherlands a height value will be available with a precision of about 5 cm, (AHN, 2000). As a test area the so-called Alblasserwaard is used, a polder of 350 km^2, directly east of Rotterdam. The location of this polder is indicated in the inset in Fig. 5. For this polder, a test data set has been kindly made available by provider

* Corresponding author.

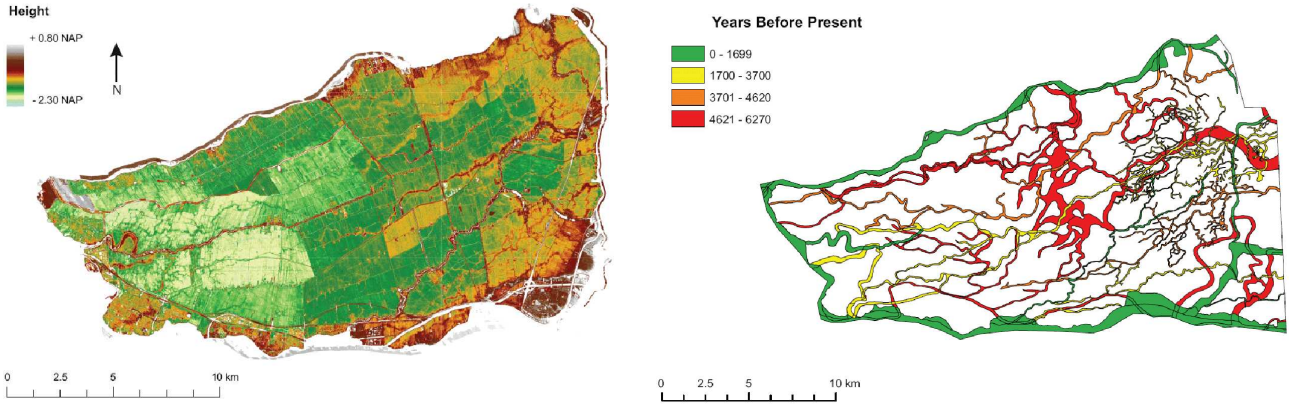

Figure 1: Alblasserwaard. **Left:** Laser altimetry data. **Right:** Paleogeographic map

Fugro Aerial Mapping B.V. and owner Waterboard Rivierenland, consisting of about 1.2 billion gridded points, see Fig. 1, left.

From this data set points representing hard infrastructure are removed in a filter procedure incorporating the Dutch topographic base map GBKN. Remaining points are classified according to four structural attributes into two classes, channel deposit and non channel deposit. In Section 2. this data filtering and classification procedure is discussed. Results are validated in Section 3. against the digital paleogeographic map and against an interpretation of drillings from the Dutch geological database DINOLoket.

2. LIDAR DATA FILTERING AND CLASSIFICATION

In this section methodology is described aiming at the classification of airborne laser altimetry points into two classes, buried channel and non-buried channel deposits. A main challenge in this research is the huge amount of input points. As the input data strongly influences the methodology, these are described first. Then it is described how laser points representing hard infrastructure are removed before describing the actual classification method.

2.1 Data description

For this research FLI-MAP400 VS laser altimetry data is used, measured by Fugro Aerial Mapping BV for the Waterboard Rivierenland. An overview of the entire data set is shown in Fig. 1, left. The data was acquired during three days in August 2007, with a minimum point density of 8 points per m^2. The absolute accuracy of a single point is reported to be 3 cm. From this raw data, Fugro derived a Digital Surface Model (DSM) by removing non-terrain points. The DSM points were consecutively resampled to a 0.5m grid using inverse squared distance weighting and organized in tiles of 1.25×1 km. In total the Alblasserwaard data set was divided into 273 of such tiles To decrease computational efforts, the .5m grid was further downsampled to a 2 m grid. As a result, the input data set for this research consists of roughly 75 million points.

2.2 Removing non-field objects

In the gridded FLI-MAP data still objects like roads, trenches, buildings and water surfaces are present. If unaddressed these objects complicate the detection of buried channel deposits. The laser data is filtered in two steps with the purpose of only keeping data representing fields. In the first step, non-field objects are removed using a mask constructed from GBKN data, in the second step remaining unwanted objects are removed, based on a local variability analysis.

GBKN mask. The 'Grootschalige Basiskaart van Nederland' (GBKN) is the Large Scale Standard Map of The Netherlands and is the most detailed and accurate digital topographical database available in the Netherlands, (GBKN, 2009). It is scale-free, but is comparable to paper maps with a scale between 1:100 and 1:5,000. The precision of a point in comparison to another point in the surrounding is better than 28 cm in suburban areas and better than 56 cm in rural areas. The GBKN has a spaghetti-structure: it only contains classified nodes and edges, for instance road sides, water edges and building contours. Therefore the GBKN map of the Alblasserwaard has to be converted to an area map, consisting of classified segments, see Fig. 2, left and middle. This area map will then be applied as a mask to remove those laser points that are in a polygonal segment from an unwanted class, like 'road'.

To create segments, the GBKN lines have to be automatically connected and converted into classified segments. However, there are errors in the database: lines sometimes do not connect exactly or lines intersect without a node. Such situations have to be identified and adapted. Lines in the GBKN that do not exactly connect are attached to the nearest line or node within a certain distance threshold in a snapping procedure. Here a threshold of 10 cm is used. Self intersections without nodes are removed by adding nodes to the intersection points. Around the resulting area mask, an additional buffer of 3m is added to further limit the influence of unwanted objects: for example, ground close to a road is often disturbed, and cannot be considered as representative for the situation in a field.

Despite this filtering method, unwanted features still remain in the LIDAR data, see Fig. 2, right, like small trenches and other objects not (yet) registered in the GBKN database. To further decrease the influence of unwanted features, isolated points and points with a high local variance were additionally removed.

2.3 Channel classification

The points remaining after the removal of non-field objects are classified into two classes by means of structural classification. For this purpose first structural attributes are determined at each grid point. As a result at each grid point a multi variate feature vector is created that can be used as input for a standard remote sensing classification method.

Slope and curvature attributes At each remaining LIDAR point, the four following attributes are determined: slope, curvature, TPI and smoothed TPI. Slope is chosen as an attribute because at both sides of a buried channel, the elevation is increasing with

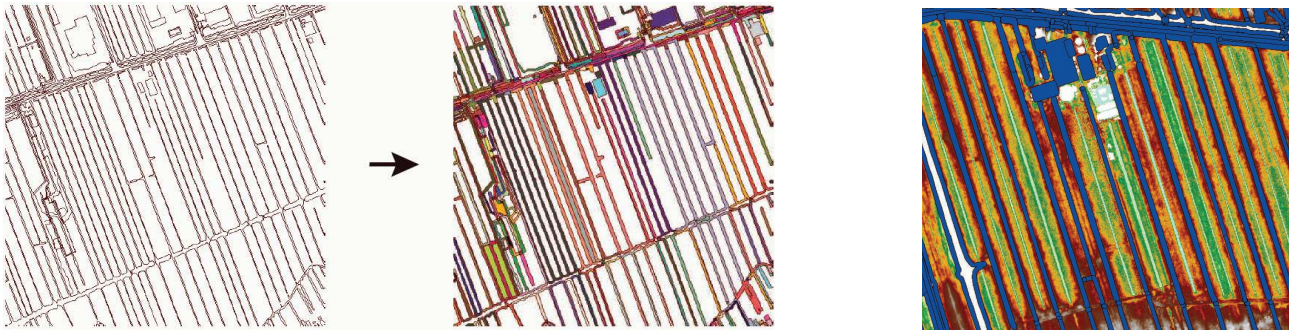

Figure 2: GBKN mask **Left:** Original GBKN line data; **Middle:** Final GBKN mask. **Right:** GBKN mask overlaid on LIDAR data.

respect to the surrounding field. To derive slope, a plane is fitted by least squares to a suited squared neighborhood of a LIDAR point. From the planar parameters, an estimation of the local slope is derived using Horn's method, (Burrough and McDonnell, 1998). Curvature is chosen as an attribute because the terrain at an elevation caused by a buried channel is convex as it is locally protruding. Flat terrain has a mean curvature of zero, while convex terrain has positive mean curvature. Here an approximation of mean curvature is derived from local partial derivatives by locally fitting a second degree polynomial surface to a suited squared neighborhood, see for more details (Besl and Jane, 1986) and (Nahib, 1990).

To reduce the computational costs of the least squares adjustment involved in the many slope and curvature determinations, a down-sampling strategy is applied. After an analysis of different down-sampling rates, in which slope values obtained from a down-sampled data set where compared to slope values from the full 2m grid input data set, it was decided to use only 10 % of the data.

TPI and smoothed TPI attributes The Topographic Position Index (TPI) is a measure of the elevation of a location compared to the surrounding landscape, (Weiss, 2001). To compute the TPI-value of a single pixel the difference between its elevation and the average elevation of a neighborhood around that cell is calculated. Most frequently an annular neighborhood is used, that is, all cells between a certain minimal and maximal distance are used in the calculation. A positive TPI-value means that the cell is higher than its surroundings (at the specified neighborhood size) while negative values mean it is lower. A TPI-value of zero indicates that the cell either lies on a flat area or on a constant slope. The TPI is of course strongly dependent on the scale. Here TPI-values are computed using a minimal distance of 80 m and a maximal distance of 100 m. From the TPI-values also a fourth attribute is determined, the smoothed TPI. This is just the mean of the TPI values in a 49 $\times$ 49 grid points window and helps to distinguish between small and large scale topographic features.

Maximum likelihood classification As a result of the structural attribute determination, at each grid point a 4D attribute vector is given, consisting of slope, mean curvature, TPI and smoothed TPI attribute values. The availability of these attribute vectors allows us to apply standard classification techniques from remote sensing. Here Maximum Likelihood classification is applied.

The Maximum Likelihood Classifier, (Gao, 2008), uses statistics from class signatures to determine if a given pixel belongs to a class. Each class signature is derived by manually selecting small areas that are known to belong to a certain class. These areas are called training samples. The training samples in this research have been selected based on manual interpretation of the height data and by looking at independent reference data, in this case the digital paleogeographic map, compare Fig. 1, right. In Fig. 3 the location of the training samples is shown. The results of the classification were slightly cleaned using the morphological operators 'majority filter' and 'conditional dilation' to remove small outlying classification results and fill small holes, e.g. (Jain, 1989).

3. RESULTS, VALIDATION AND DISCUSSION

In this section the results of the automatic classification of the Alblasserwaard LIDAR data are presented, validated and discussed. First visual results are discussed. Then two validation methods based on soil drillings are described, together with the results of the actual validations.

3.1 Visual validation

The blue points in Fig. 5 indicate LIDAR 2m grid points classified as 'buried channel'. Clearly some more or less connected channel structures in East-West direction are recognizable. Simultaneously, many thicker fragments classified as 'buried channel' are visible. Based on a visual evaluation it is not directly obvious if these thicker fragments indeed correspond to channel deposits. Also anomalies are visible in the classification results: thin, straight lines appear at many locations and are mainly corresponding to terrain close to roads and ditches.

Fig. 8 shows a zoom-in of the classification results, again in blue, superimposed on areal imagery data. The area in Fig. 8 approximately corresponds to the red rectangle in Fig. 5. This image

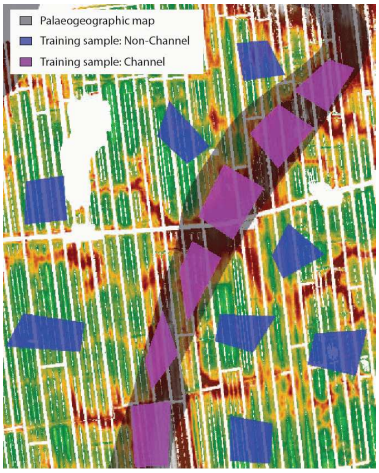

Figure 3: Training samples used for the classification process.

confirms that the classification is still influenced by infrastructure: The classification algorithm reports buried channels near and at farmyards and along a small ditch which indicates that the GBKN infrastructure database is not complete and that the filtering procedure should be further improved.

3.2 Drilling data description

For this research two independent validation data sets based on soil drillings are available. The first is a map product, the second set consist of a large amount of single drillings, interpreted by the authors.

Digital paleogeographic map A digital paleogeographic map of the complete Rhine-Meuse delta during the Holocene (including the locations of buried channel deposits) is described in (Berendsen and Stouthamer, 2001). The Alblasserwaard section of this map, Fig. 1, right, is used as validation in this research. 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. The map is stored in vector format, each individual area consists of a polygon. For each area up to 12 different attributes are stored like channel size, channel length, age, year of beginning, year of ending, etc. The age of the channels on the map are given in years Before Present, where Present is defined as the year 1950. For this research four main age categories are distinguished, indicated by different colors in Fig. 1. These periods have been manually chosen based on the distribution and amount of channels abandoned in these periods.

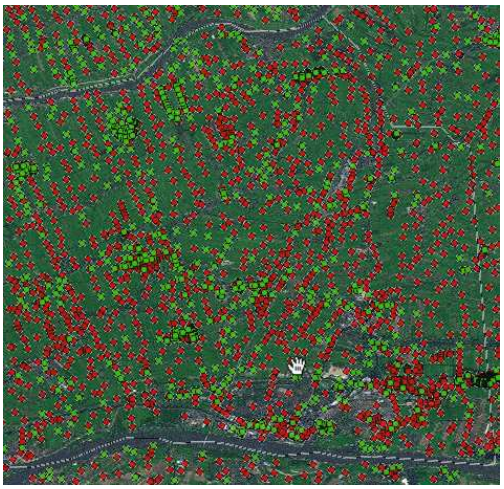

Figure 4: Classified DINO drillings.

DINO drillings The DINO database contains data and information of the subsurface of The Netherlands, (DINOLoket, 2000). The archive contains among others shallow boring measurements that are suitable to use as reference data for this research. They cover primarily the shallow subsurface and contain standardized information about the type of sediments and their depth. In total 2 680 individual drillings were available for the Alblasserwaard. The eastern part has a high drilling density, in the western part only a very limited number of drillings is available.

After importing 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 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 channel. The reason to discard the top 3 meter is that sand layers can be present there due to other reasons, like construction works. The analysis of all of the drillings in the eastern part is shown in Fig. 4. This form of automatic interpretation of drilling data is prone to errors. This means that in this case the amount of correctly interpreted drillings is largely unknown. Still in Fig. 4 the spatial correlation between drillings and LIDAR classification results is visible.

3.3 Validation results

Further visual validation is obtained by comparing the automatically classified LIDAR points to the digital paleogeographic map and to the classified DINO drillings. For this project this was done using the ESRI Flex viewer, (ESRI, 2009). This program allows internet users to simultaneously view within their normal browser the different spatial layers on available background imagery, just as within a GIS environment. A screen shot is shown in Fig. 8. According to the digital paleogeographic map this figure contains buried channel deposits from three periods, compare Fig. 1, right. In red some relative old (6270-4621 yBP) and wide buried channels are visible, in orange another wide, slightly younger (4620-3701 yBP) channel is given. while some relatively young (3700-1700 yBP) smaller channel deposits are shown in yellow. The LIDAR data classified as channel deposit (in blue) gives the best match with the orange channel, while some matching results on the yellow channels are found as well. There seems to be hardly no correlation between the blue LIDAR channel deposits, and the large and old red channels. Similarly the classified DINO drillings give good agreement over the orange channel, while the DINO drillings give mixed responses over the red channel. Many DINO drillings outside the areas classified by the two other methods as buried channel are indeed red, but also here exceptions exist.

Table 6: Classified LIDAR vs. paleogeographic map

	Map with all channels	
LIDAR channel	3.5 %	6.3 %
non-channel	23.5 %	66.7 %
	Map 4620-1700 yBP	
LIDAR channel	2.0 %	7.6 %
non-channel	5.4 %	84.9 %
	channel	**non-channel**

These observations are partly confirmed by the numeric comparison over the region of Alblasserwaard as a whole. In Table 6, two confusion matrices of the LIDAR buried channel classification compared to the paleogeographic map are given. The top matrix compares the LIDAR classification to all buried channels shown in Fig. 1, right; in the bottom matrix the comparison is restricted to those channels in the paleogeographic map that are dated between 4620 and 1700 yBP, i.e. the yellow and orange channels in Fig. 1, right. The diagonals give the percentages of pixels where map and classified LIDAR agree, the upper right entry is the percentage of pixels that are non-channel in the reference map, but are classified as channel. The reverse holds for the bottom left entry. Although the overall classification accuracy, i.e. the trace of the confusion matrix, equals 70 %, kappa, a measure of similarity without chance agreement, only equals $\kappa = 0.06$. By removing the youngest and oldest channel class from the comparison, the amount of agreement improves to 87 % with an associated kappa value of $\kappa = 0.21$.

There are several possible reasons for this lack of agreement. First of all it should be noted that a condition for a high degree of agreement is that a buried channel deposit always results in a locally higher elevation. Although there is strong evidence that

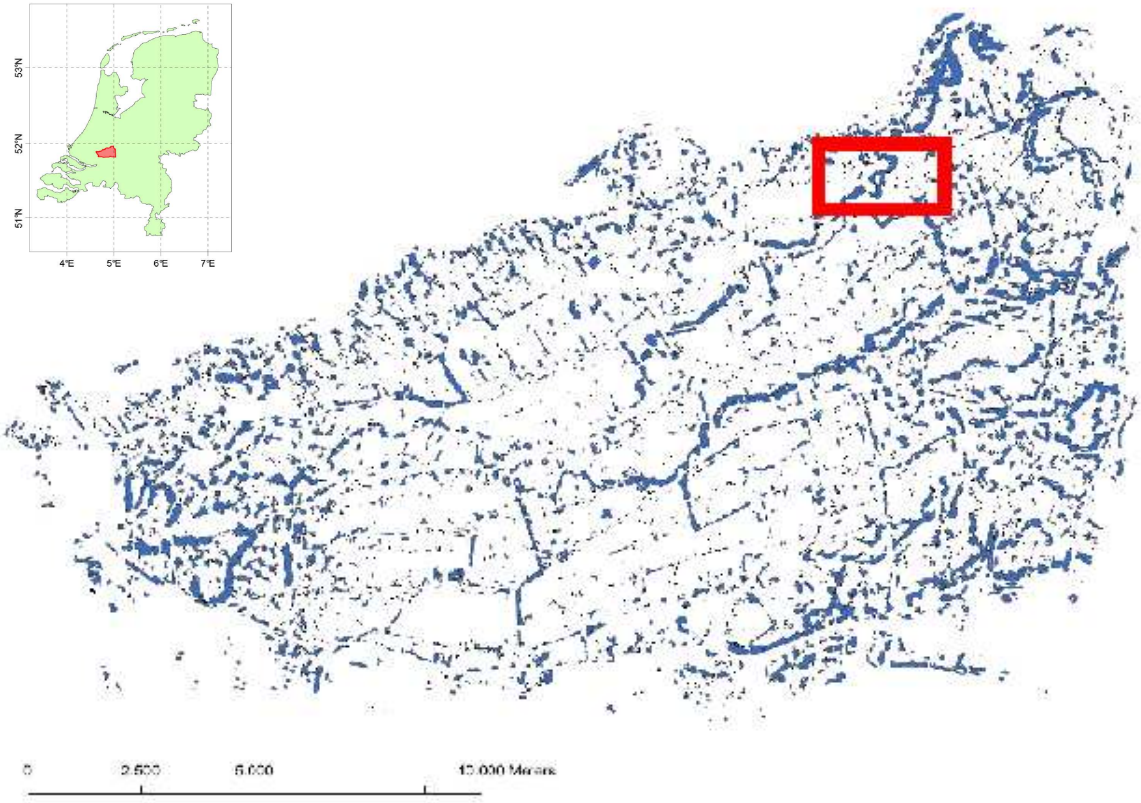

Figure 5: Buried channel classification results. The red rectangle approximately corresponds to the area of Fig. 8. The inset shows the location of the Alblasserwaard in The Netherlands.

this condition in general is fulfilled, it is not yet sufficiently clear, what buried channel characteristics result in what amount of local elevation setup. Other reasons originate in the processing of the available information. In the composition of the paleogeographic map, errors are associated to the interpretation and interpolation of the used drillings. The reason that the youngest, green, channels in the Paleogeographic map do not give a good comparison with the classified LIDAR data is simply that these channels either still exist at approximately the same location or that buildings and roads are present along or on the remains. In both cases the LIDAR data for these regions is simply filtered out in the data processing procedure.

Table 7: Classified DINO drillings vs. Classified LIDAR and paleogeographic map.

	DINO drillings	
LIDAR channel	6.3 %	7.3 %
non-channel	26.8 %	59.6 %
Map channel	15.0 %	18.6 %
non-channel	18.1 %	48.2 %
	channel	**non-channel**

In Table 7 also the confusion matrices between the classified DINO drillings and the classified LIDAR results, top, and the paleogeographic map, bottom, are given. Both the LIDAR result and the map have a comparable percentage of agreement (trace of both matrices) with the classified DINO drillings. They do however differ in the type of misclassification: in the LIDAR classification a relative large percentage of points were classified as non-channel that were channels according to our automatic interpretation of the DINO drillings. Again this could be caused by currently present infrastructure: many DINO drillings were obtained in the green zones, i.e. regions marked as young, channel deposits in the paleogeographic map, Fig. 1, right, where no reliable LIDAR surface height data is available.

4. CONCLUSIONS AND RECOMMENDATIONS

In this work, an original approach for the detection of buried channel deposits from high resolution LIDAR data has been described and validated. The first results indicate that to some extend it is possible to automatically determine the location of sand-rich channel areas: relatively young (4620-1700 yBP) and wide ($\sim$ 100m) channel deposits are often detected by the described method based on classification of a feature vector consisting of structural attributes derived from LIDAR data. Current results are however still far from a form where they could be applied in for example road construction. The results of this large case study also demonstrate that there are many assumptions/steps involved in both deriving the initial classification results and in validating these results. In future work, the impact of these assumptions on the final results should be further investigated.

Comparison to the paleogeographic map indicates that the depth, age and probably also size of the channel deposits are parameters whose influence on the relative elevation should be further investigated. The digital paleogeographic map is derived based on an interpretation of actual soil drillings and an interpolation step to connect identified channel locations to a braided network of channels. This last step has not been implemented yet for our automatic buried channel classification. To improve computational feasibility, the original LIDAR data has been downsampled in this study. To improve classification results it is recommended to start by analyzing a small area at full resolution in order to, first,

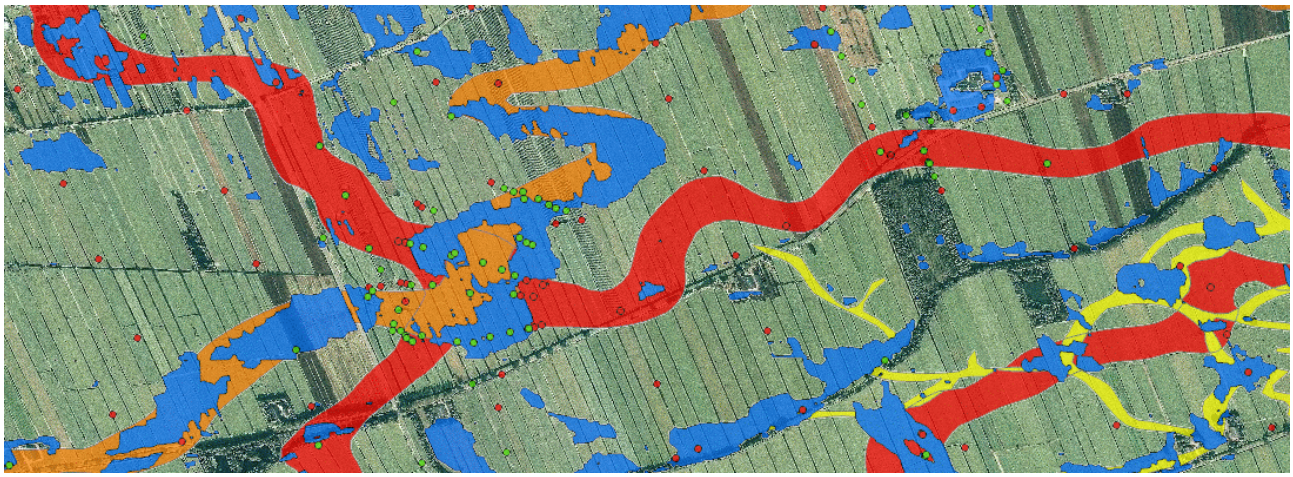

Figure 8: Buried channel classification results. **Blue:** Automatic classification LIDAR data; **Red, orange, yellow:** classification according to digital paleogeographic map, compare Fig. 1; **Red dots:** DINO drillings classified as non-channel; **Green dots:** DINO drillings classified as channels.

obtain better insight in what (channel deposit) signals are exactly present in the data and, second, to adapt the classification strategy accordingly.

REFERENCES

AHN, 2000. Productspecificaties AHN1 en AHN2. Technical report, Rijkswaterstaat, DID. http://www.ahn.nl, last visited: November 20, 2009.

Allen, J., 1965. A review of the origin and charactersitics of recent alluvial sediments. Sedimentology 5, pp. 89–101.

Berendsen, H. and Stouthamer, E., 2001. Palaeogeographic development of the Rhine-Meuse delta, The Netherlands. Koninklijke Van Gorcum.

Berendsen, H. J. A. and Volleberg, K., 2007. 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), pp. 15–22.

Besl, P. and Jane, R., 1986. Invariant surface characteristics for 3d object recognition in range images. Computer Vision, Graphics, and Image Processing archive 33(1), pp. 33–80.

Burrough, P. and McDonnell, R., 1998. Principles of Geographical Information System. Oxford University Press.

DINOLoket, 2000. Data and Information of the Subsurface of The Netherlands. Technical report, TNO. http://www.dinoloket.nl/en/DINOLoket.html, last visited: November 20, 2009.

ESRI, 2009. Get started with the sample flex viewer. Technical report, ESRI. http://www.esri.com/events/seminars/webmaps/pdfs/handout.pdf, Accessed: November 20, 2009.

Gao, J., 2008. Digital Analysis of Remotely Sensed Imagery. McGraw-Hill Professional.

GBKN, 2009. Grootschalige BasisKaart van Nederland. http://www.gbkn.nl/nieuwesite/html/engelsesite.html, Accessed: November 20, 2009.

Humme, A., Lindenbergh, R. and Sueur, C., 2006. Revealing celtic fields from lidar data using kriging based filtering. In: Proceedings ISPRS Commission V Symposium, 'Image engineering and vision metrology, Dresden.

Jain, A., 1989. Fundamentals of digital image processing. Prentice-Hall, Inc., New Jersey.

Kakiuchi, T. and Chikatsu, H., 2008. Robust extraction of ancient burial mounds in brushland from laser scanning data. IAPRS XXXVII(B5), pp. 341–346.

Kooistra, M. and Maas, G., 2008. The widespread occurrence of Celtic field systems in the central part of the Netherlands. Journal of Archaeological Science 35(8), pp. 2318–2328.

M. Doneus, C. Briese, M. F. and Jammer, M., 2008. Archaeological prospection of forested areas using full-waveform airborne laser scanning. Journal of archaeological science 35(4), pp. 882–893.

Munstermann, W., Ngan-Tillard, D. and Venmans, A., 2008. Total engineering geological approach to motorway construction on soft soils. In: Proceedings EurEnGeo conference, Madrid, Spain, pp. 1–10.

Nahib, N., 1990. Algebraic error analysis for surface curvature and segmentation of 3-d range images. Pattern Recognition 23(8), pp. 807–817.

Perron, J., Kirchner, J. and Dietrich, W., 2008. Spectral signatures of characteristic spatial scales and non-fractal structure in landscapes. Journal of Geophysical Research - Earth Surface 113(F04003), pp. 1–14.

Weiss, A., 2001. Topographic position and landforms analysis - poster presentation. In: ESRI User Conference, San Diego, United States, p. 1.

ACKNOWLEDGMENTS

The authors would like to thank Waterboard Rivierenland and Fugro Aerial Mapping B.V. for providing them with the airborne laser data. Esther Stouthamer from Utrecht University is thanked for providing the authors with a digital version of the paleogeographic map.