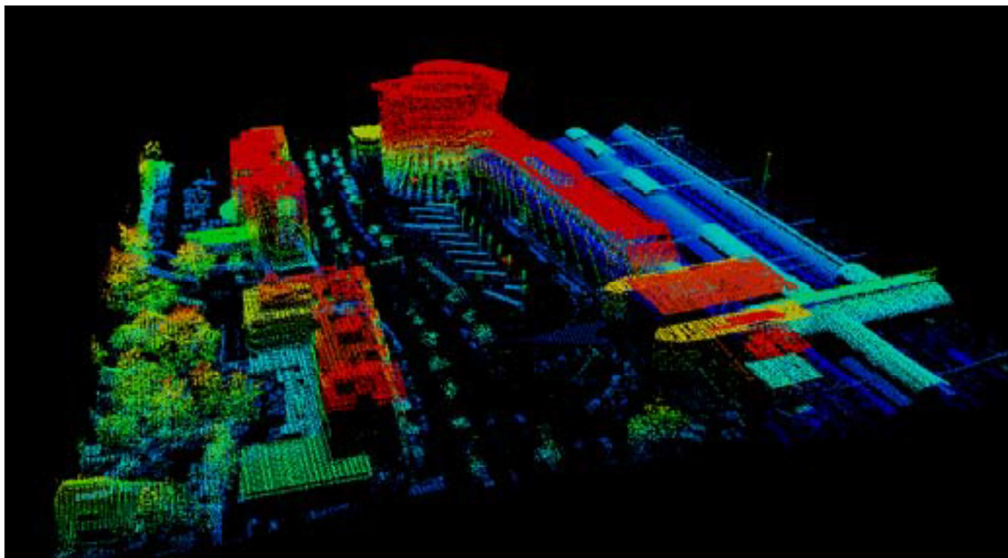


User requirements Massive Point Clouds for eSciences (WP1)



Author: P.M. Suijker (TU Delft Library)
Co-authors: I. Alkemade (Rijkswaterstaat Centrale Informatievoorziening)
M. P. Kodde (Fugro GeoServices B.V.)
A.E. Nonhebel (TU Delft Library)

Date: 2014, April 25

Version: Final version, plus some small adaptations

1 Table of Contents

2	Introduction	4
3	Interview/questionnaire	4
4	Current use of raster and point cloud data.....	5
4.1	Shipment of AHN data	5
4.2	The use of AHN raster data and AHN point cloud data	6
4.2.1	AHN raster data	6
4.2.2	AHN point cloud data	7
4.3	Research examples	8
4.3.1	Modeling pluvial flooding damage	8
4.3.2	The Polder Atlas of the Netherlands	8
4.3.3	Urban heat islands.....	8
4.3.4	Measure cubic contents	9
4.3.5	Urban scene reconstruction	9
4.4	The use of software packages.....	9
5	User requirements for the point cloud spatial DBMS	10
5.1	Selections	10
5.1.1	Selection: rectangle.....	10
5.1.2	Selection: line + buffer	10
5.1.3	Selection: polygon	11
5.1.4	Selection: attributes	11
5.1.5	Selection: search terms	11
5.2	Calculations.....	12
5.2.1	Calculations: point clouds density.....	12
5.2.2	Calculations: slope orientation, steepness.....	12
5.2.3	Calculations: statistic values.....	13
5.2.4	Calculations: plane detection	13
5.2.5	Calculations: point normal estimation	14
5.3	Analyses	14
5.3.1	Analysis: spatial analyses with other tables	14
5.3.2	Analysis: temporal differences.....	15
5.3.3	Analysis: pattern.....	15
5.4	Access.....	16
5.5	View	16

5.6	Navigate	17
5.7	Download form and format	18
5.8	Attributes	18
5.9	Co-ordinate systems	19
6	Conclusions	20
7	References	21
7.1	Publications.....	21
7.2	Websites	21

2 Introduction

This report is a milestone in work package 1 (WP1) of the project Massive point clouds for eSciences. In WP1 the basic functionalities needed for a new Point Cloud Spatial Database Management System are identified. This is achieved by (1) literature research, (2) discussions with the project members, and (3) customer surveys. In this work package three organizations TU Delft, Rijkswaterstaat and Fugro have worked together to assemble the user requirements for a new Point Cloud Spatial Database Management System. The following persons were involved in this work package.

A.E. Nonhebel (Ardi)	TU Delft, TU Delft Library/3TU.Datacentrum (work package manager)
Ir. I. Alkemade (Ingrid)	Rijkswaterstaat, Centrale Informatievoorziening
Ir. M.P. Kodde (Martin)	Fugro GeoServices B.V.
Drs. T.P.M. Tijssen (Theo)	TU Delft, Onderzoeksinstituut OTB, department of GIS Technology
Prof.dr.ir. P.J.M. Oosterom (Peter)	TU Delft, Onderzoeksinstituut OTB, department of GIS Technology
Drs. P.M. Suijker (Paul)	TU Delft, TU Delft Library, section Research support

During the investigation the Up-to-date Height Model of the Netherlands¹ (AHN) datasets have been used as an example of the type of data that will be included in the new point cloud spatial DBMS. The AHN data consists of raster data and point cloud data. Both data types have been considered.

3 Interview/questionnaire

In preparation to this report we have had a lot of support from researchers and professionals who have used the Up-to-date Height Model of the Netherlands (AHN) data in their research or in their daily work. We would like to thank the following persons that have participated in drawing up these user requirements, either by means of an interview or by filling in the questionnaire.

Ir. S. Gaitan	TU Delft researcher, faculty of Civil Engineering and Geosciences, department of Water Management	interview
M.T. Pouderoijen	TU Delft researcher, faculty of Architecture, department of Landscape Architecture	interview
Dipl.ing. A. Wandl	TU Delft researcher, faculty of Architecture, department of Urbanism	interview
Dr. ir. O.A.C. Hoes	TU Delft researcher, faculty of Civil Engineering and Geosciences, department of Water Management	poll
Drs. M.J. Zeylmans van Emmichoven	Utrecht University researcher, faculty Geosciences	interview
Dr. K.M. Cohen	Utrecht University researcher, faculty Geosciences	interview
Ir. F.J. van Leijen	TU Delft researcher, faculty of Civil Engineering and Geosciences, department of Geoscience and Remote Sensing	interview
Drs. G. Jimmink	Rijkswaterstaat, GIS-coordinator, Grote Projecten en Onderhoud (GPO)	poll
R. Schrijver	Rijkswaterstaat, GIS-analist, Water, Verkeer en Leefomgeving (WVL)	poll
Drs. Q.J. Lodder	Rijkswaterstaat, senior advisor, Water, Verkeer en Leefomgeving (WVL).	poll

¹ Actueel Hoogtebestand Nederland (AHN)

	TU Delft researcher, faculty of Civil Engineering and Geosciences, Coastal Engineering	
Drs. H.S. van der Wal	Rijkswaterstaat, GIS specialist, Centrale Informatievoorziening (CIV)	poll
Ing. P. Cornelissen	Rijkswaterstaat, Water, Verkeer en Leefomgeving (WVL)	poll
Ing. H. Faber	Rijkswaterstaat, consultant dike and dam safety and water systems development	poll
A. D. Grul	Rijkswaterstaat, Centrale Informatievoorziening (CIV)	poll
Ir. D.J. ten Napel	Rijkswaterstaat, Centrale Informatievoorziening (CIV)	poll
Ing. P. Bakhuizen	Fugro GeoServices B.V.	poll
Ir. J. Haan	Fugro GeoServices B.V.	poll

4 Current use of raster and point cloud data

4.1 Shipment of AHN data

One of the reasons for building a new point cloud spatial DBMS is to improve the dissemination of point cloud data, like AHN datasets. We asked the researchers to indicate how they receive their AHN raster data and point cloud data now.

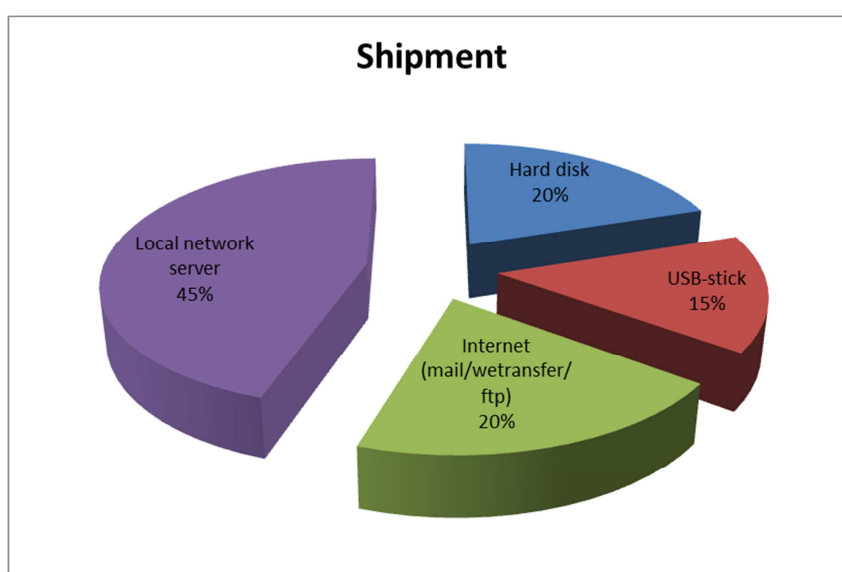


Figure 1 – Distribution of AHN data

There is a large difference between the distribution of data between organizations. The TU Delft mainly copies fractions of the datasets to external devices like hard disks and USB-sticks. Rijkswaterstaat has stored the data on a local network server that can be accessed by most employees.

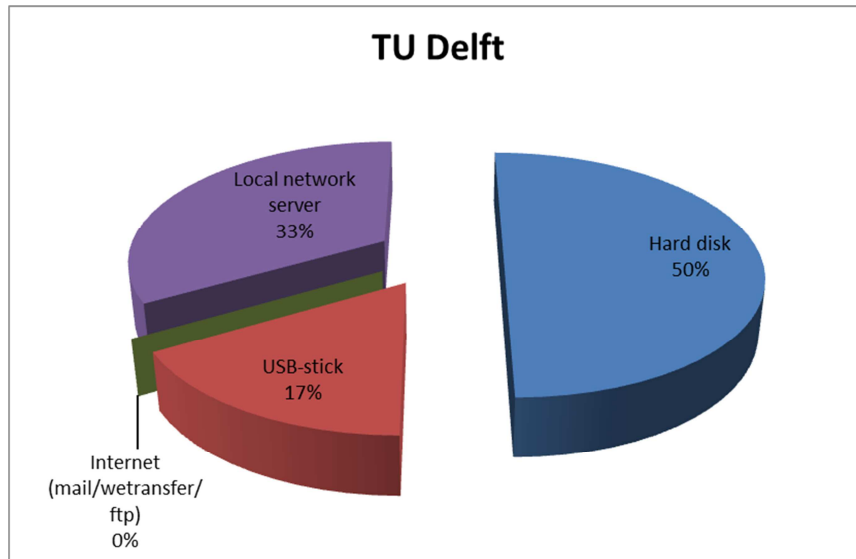


Figure 2 - Distribution of AHN data at TU Delft

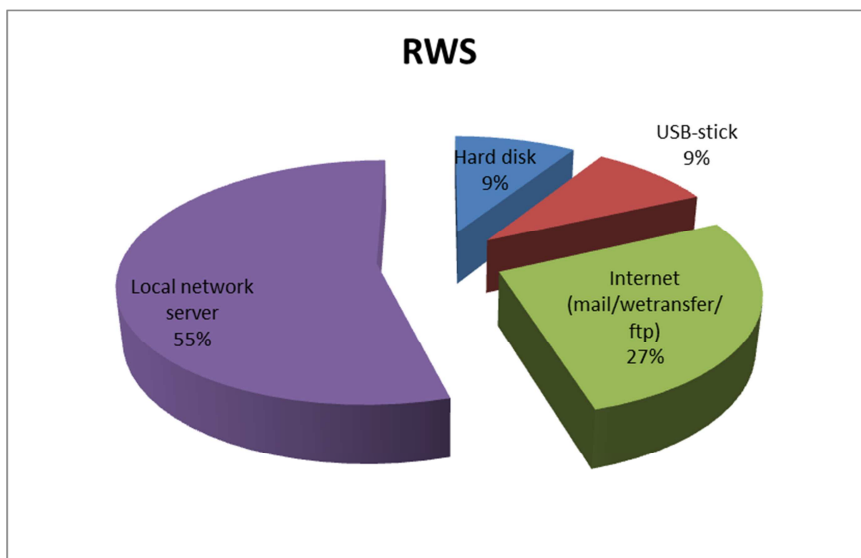


Figure 3 - Distribution of AHN data at Rijkswaterstaat (RWS)

The staff members and students from Utrecht University have access to a network server where the AHN raster data is stored in a database. Students have to fill out an application after which they can access the database with their NetId. Utrecht University doesn't have any point cloud data available yet for their users. Fugro is somewhere in between: one third of the data is distributed via external devices, one third via internet and one third via local network servers.

4.2 The use of AHN raster data and AHN point cloud data

4.2.1 AHN raster data

The use of AHN1 and AHN2 raster data is widespread in all organizations that took part in this survey. The AHN1 five meter raster data (filtered) and the AHN2 half meter raster (filtered with interpolation) are the datasets that have been used most commonly. Staff members from Rijkswaterstaat mentioned that the AHN2 five meter raster data fits perfectly to their bathometric (depth) data.

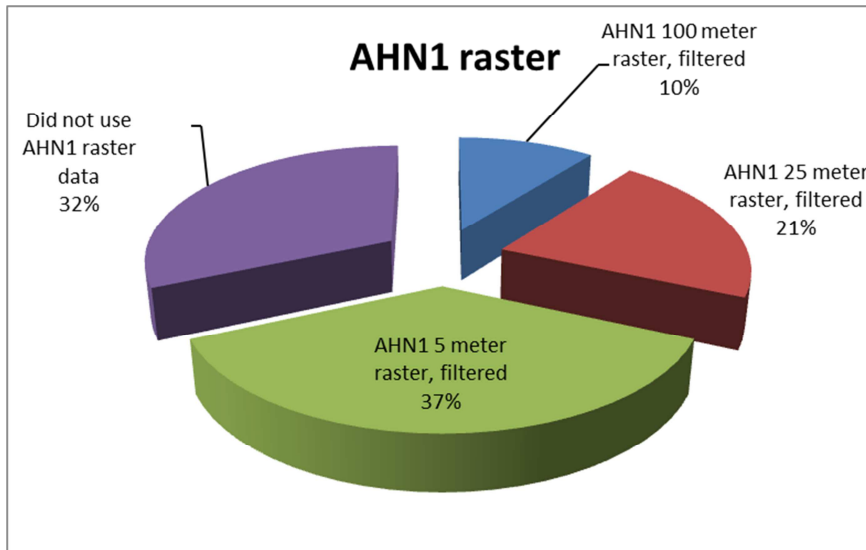


Figure 4 - Use of AHN1 raster data

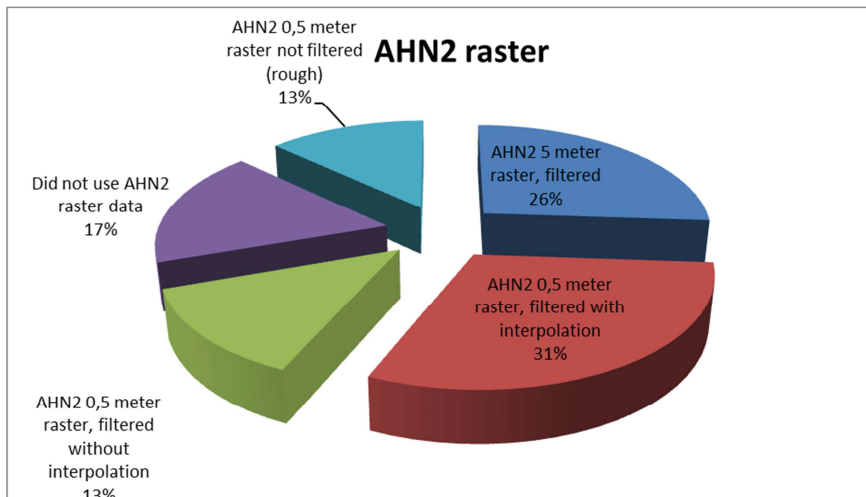


Figure 5 - Use of AHN2 raster data

4.2.2 AHN point cloud data

The AHN point cloud data have not been used as extensively as the AHN raster data. Most respondents contend that the raster data contain enough information to perform their research or task. At the same time they claim that they would use the point cloud data if the datasets were more readily available or if they could benefit more easily from the additional information that is available in the point cloud data (e.g. in the form of attributes). Researchers at TU Delft and Utrecht University have been using point cloud data, either to gauge the potential of the point cloud data or in their current research [Straatsma, 2008]. Staff members from Rijkswaterstaat and Fugro have used point cloud data more often than researchers at the universities.

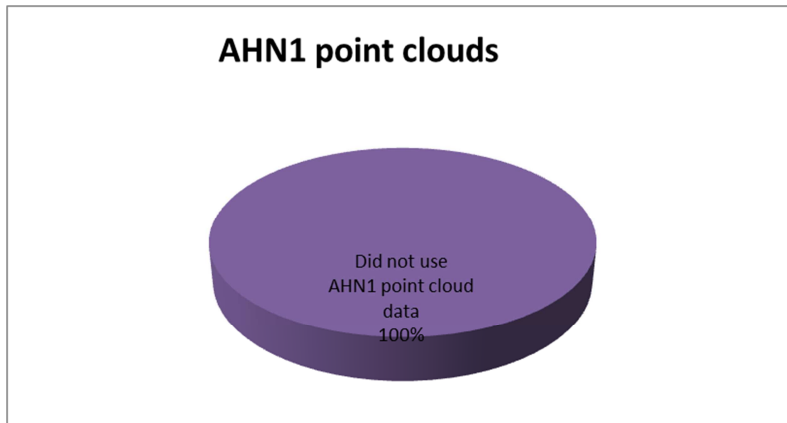


Figure 6 - Use of AHN1 point cloud data

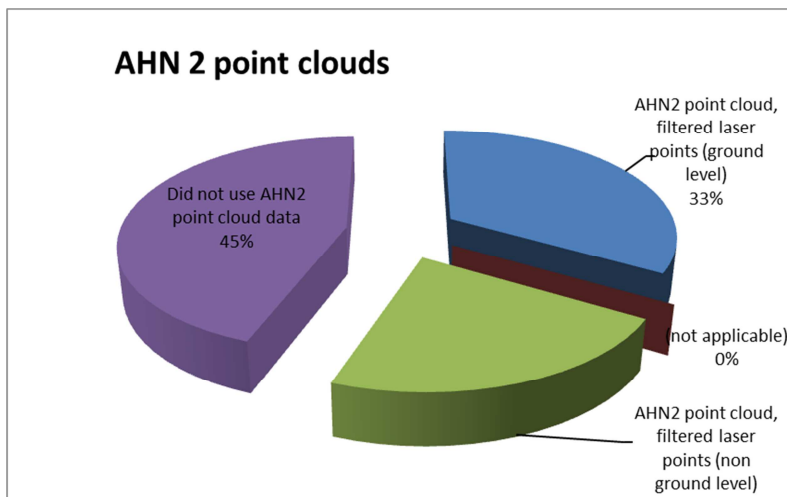


Figure 7 - Use of AHN2 point cloud data

4.3 Research examples

A few examples have been added to give a picture of the kind of research that has been executed with raster and point cloud data.

4.3.1 Modeling pluvial flooding damage

Santiago Gaitan is using AHN2 0,5 meter raster data for his doctoral research [Gaitan, 2013]. He is using the AHN raster data to assess the urban vulnerability to pluvial flooding in de city of Rotterdam. Gaitan uses a hydrological analysis tool (D8) to determine the flow direction of the rainwater [Website: Whitebox Geospatial Analysis Tools]. In the near future he will also use TIN data structures for his models [Website: Triangulated irregular network (TIN)]. See also [Ledoux, 2009].

4.3.2 The Polder Atlas of the Netherlands

Steffen Nijhuis and Michiel Pouderoijen have used AHN 5 meter raster to produce shaded relief maps of Dutch polders [Steenbergen, 2009]. He has used AHN2 point cloud data to classify buildings and landscape elements. This research has not resulted yet in a publication.

4.3.3 Urban heat islands

Alexander Wandl has used AHN raster data to calculate the average building volume and shade factor for Amsterdam in his urban heat island research [Hoeven, 2013]. In the near future he would like to use AHN2 point cloud data, using the attributes to classify roof types and other objects.

4.3.4 Measure cubic contents

Gerben Jimmink uses AHN raster data, among others, to measure the cubic content of contaminated soil in redevelopment areas. He also used a combination of the Basisregistraties Adressen en Gebouwen (BAG) and AHN2 raster data for building height measurements in sound modeling.

4.3.5 Urban scene reconstruction

Some more examples of actual use of point clouds in science (not interviewed by us):

1. Researchers from Utrecht University have studied laser point data to identify rectangles for urban scene reconstruction [Lankveld, 2011]
2. Researchers from the University of Innsbruck have used point cloud data in landslide modeling studies [Rutzinger, 2013]
3. At the SilviLaser conference in Vancouver 2012, researchers presented a number of studies based on point cloud data [Website: SilviLaser conference]

4.4 The use of software packages

Most respondents use the ArcView and ArcGIS desktop software or the Fugro Viewer to view the AHN data. A Fugro staff member also mentioned QINSy and Qloud, hydrographic data acquisition, navigation and processing software packages from QPS. The researchers at the universities also mentioned GRASS GIS. The functionality of GRASS matches the ArcGIS functionality but the software package is difficult to master and only recently available on other than Linux platforms. In spite of the popularity of ArcGIS, the majority of the researchers prefer to process the data with software they developed themselves using programming language like Python, MATLAB, R and their respective packages for manipulation and analysis of spatial information. They argue that they want to have absolute control about the processing of the data. Some TU Delft researchers have developed their own interpolation tools, because they were not satisfied with the standard interpolation offered by AHN. Some other tools were mentioned: LAsTools, sky-view factor tools and the Fugro Lidar processing tool.

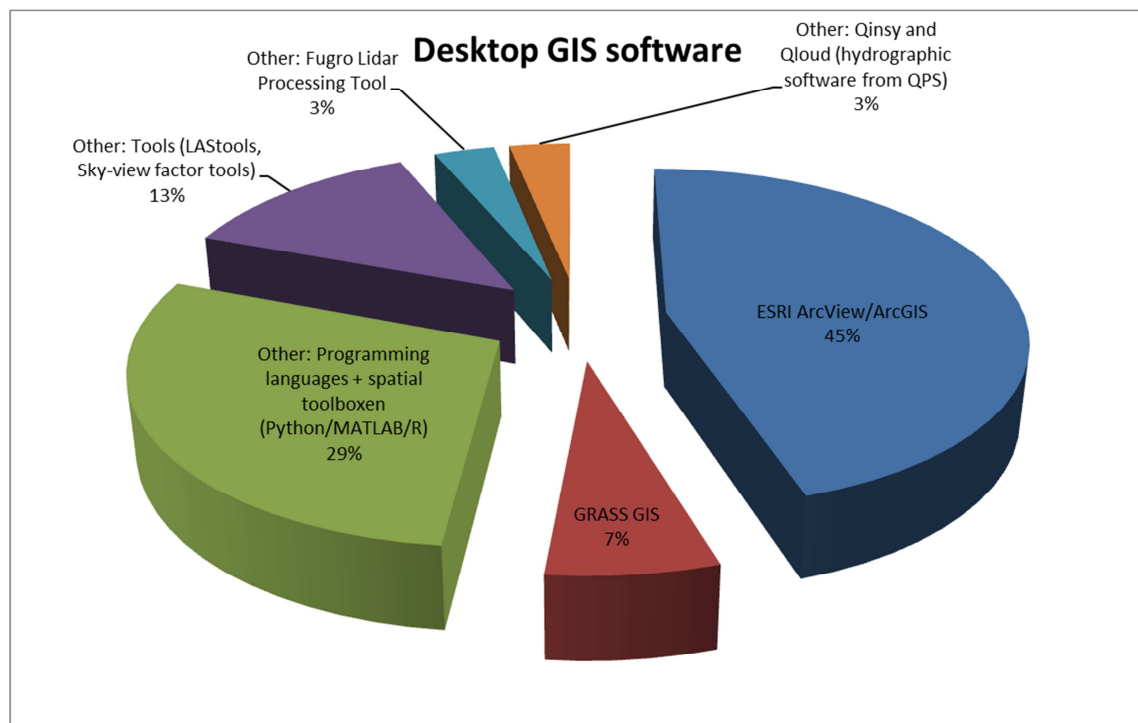


Figure 8 - Use of desktop GIS software

5 User requirements for the point cloud spatial DBMS

The user requirements have been listed in three categories: selections, calculations and analyses.

The user requirements have been listed in a number of categories: functionality within the system (selections, calculations and analyses), presentation functionality (access, view, navigate) and download functionality. Some functionality was examined in more detail (attributes). In order to prioritize the individual items the MoSCoW-method is used [Website: MoSCoW].

5.1 Selections

A number of options for selections were presented. One researcher preferred to select the data based upon the current AHN map sheet division (e.g. U37bz1_25).

5.1.1 Selection: rectangle

The selection of points based on a rectangle is an absolute must for the new point cloud spatial DBMS. One researcher noted that if he was able to perform an polygon selection, there is no explicit need for a rectangle search, and thus he didn't tick the option rectangle.

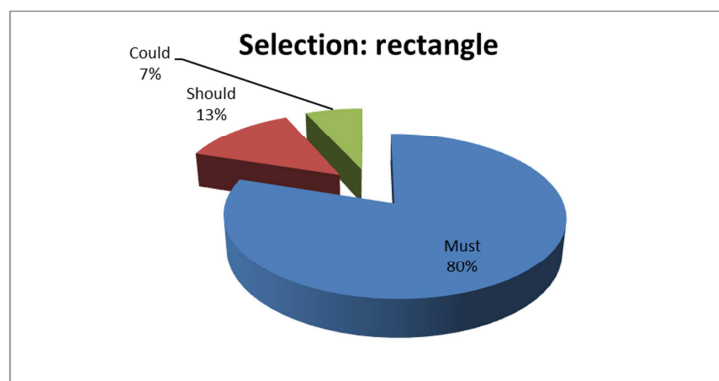


Figure 9 - Selection of points based upon a rectangle search

5.1.2 Selection: line + buffer

The selection of points based upon a line and a corresponding buffer zone, e.g. a river including 100 meters on both side of the river, is a desired selection method. Especially the researchers in water management would like to be able to use this functionality.

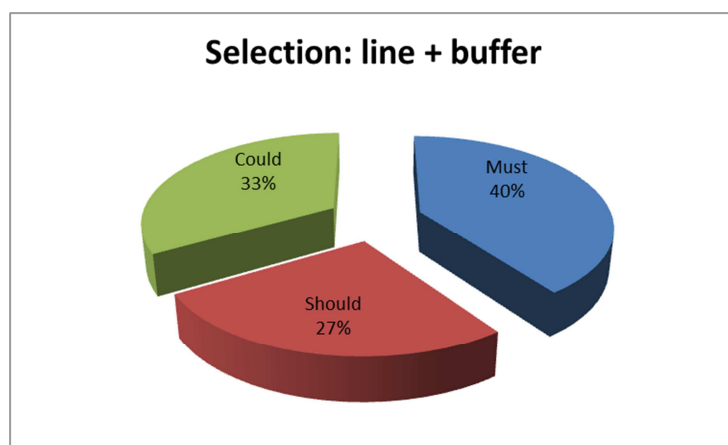


Figure 10 - Selection of points based upon a line and buffer search

5.1.3 Selection: polygon

The selection of points based upon a polygon is much sought-after. Researchers want to be able to select points from a river basin or a municipal boundary.

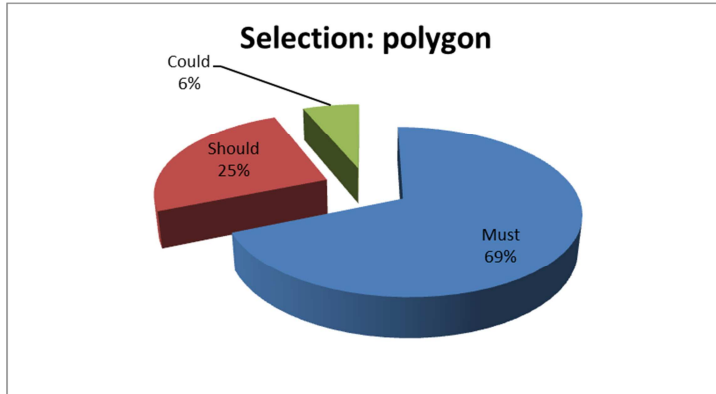


Figure 11 - Selection of points based upon a polygon search

5.1.4 Selection: attributes

Most researchers would like to be able to select points based upon attribute search. For example they would like to be able to select on RGB (red, green and blue) values or the intensity of the pulse. If possible they would also like to select the points based on object categories (for example all buildings with a flat roof). A necessary condition for this type of search is a pre-categorization of the existing point cloud data, (partially) based on attributes.

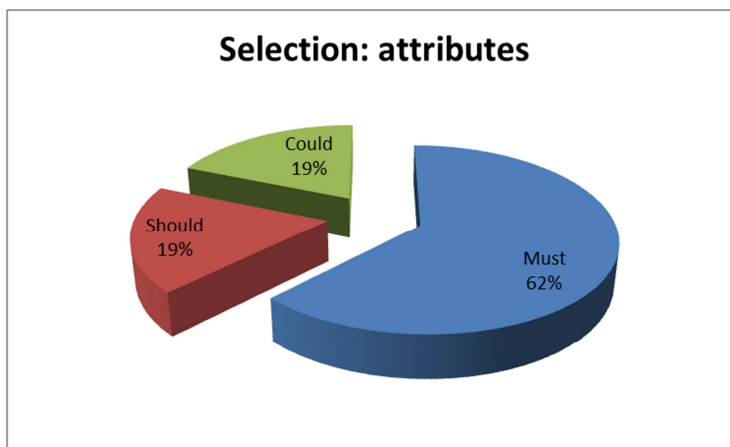


Figure 12 - Selection of points based upon an attribute search

5.1.5 Selection: search terms

The selection of points based upon search terms (e.g. Amsterdam or de Zaanstreek) is not very popular.

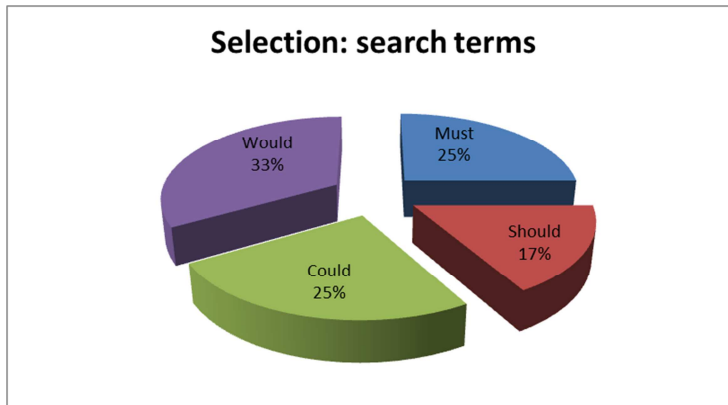


Figure 13 - Selection of points based upon search terms

5.2 Calculations

We asked the researchers what kind of calculations they would like to be able to do within the new point clouds spatial DBMS. Many researchers remarked that they prefer to make calculations using their own software. If possible they would like to be able to preview the results of calculations within the system before downloading the points to their local system and doing their own re-calculations. The most important calculations that the researchers would like to be able to make are: (1) calculation of statistical values, like the average height of an area, lowest point, highest point (2) the (mean) slope orientation and angle of inclination of an area (3) points density. Plane detection and point normal estimation were less sought-after. Researchers from RWS also want to be able to create profiles and calculate cubic contents of volumes.

5.2.1 Calculations: point clouds density

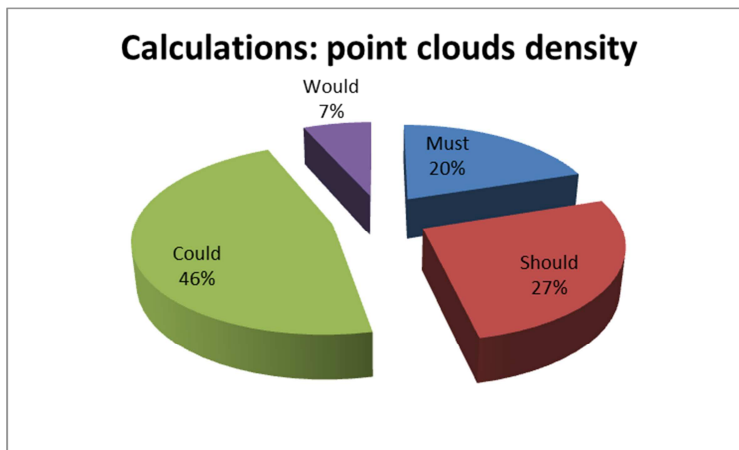


Figure 14 - Calculations of point cloud density

5.2.2 Calculations: slope orientation, steepness

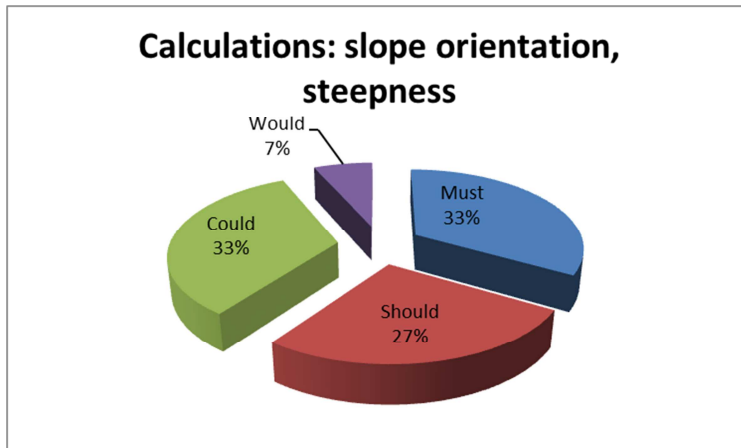


Figure 15 - Calculations of slope orientation and slope steepness

5.2.3 Calculations: statistic values

Determining statistical values for an area, like average height values, highest and lowest height values is in high demand. One respondent noted that he would rather have statistical values for a point, rather than an area.

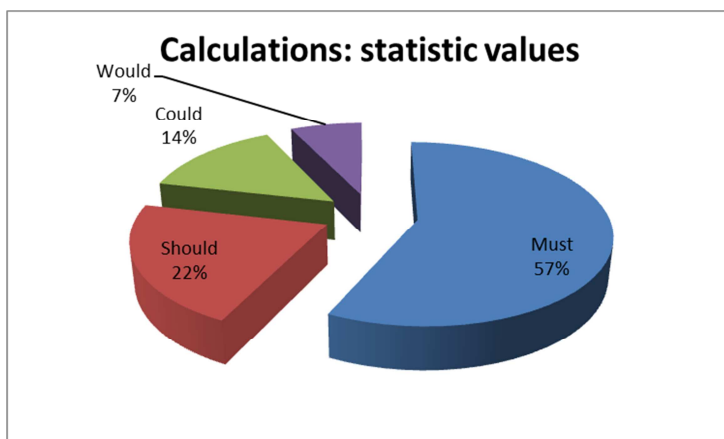


Figure 16 - calculations of statistic values

5.2.4 Calculations: plane detection

Detect mathematical planes from a point cloud. This technique can be used to detect multiple planes from a 3D point cloud, like the roof and walls of a house.

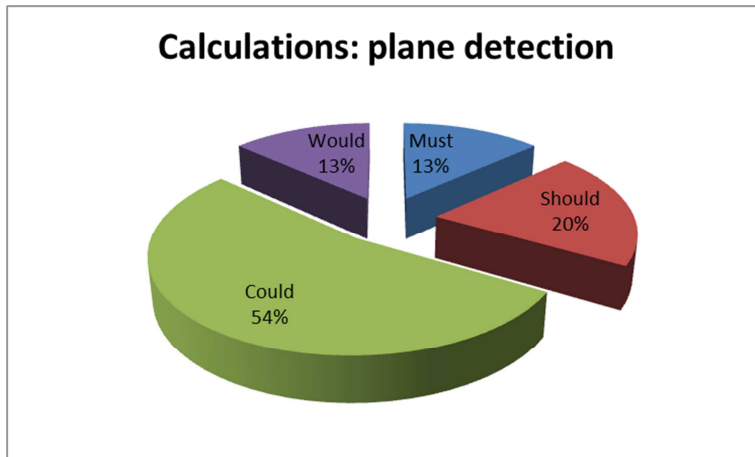


Figure 17 - Calculations of plane detection

5.2.5 Calculations: point normal estimation

Estimate normals (standard planes) on unorganized point clouds.

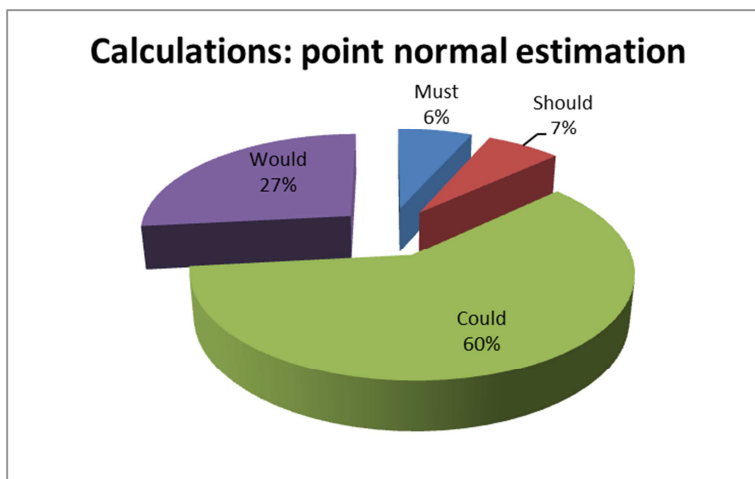


Figure 18 - Calculations of point normal estimation

5.3 Analyses

Researchers would like to (1) compare point clouds with data coming from other sources, and they would like to be able to (2) compare point clouds with older point clouds of the same area. A researcher remarked that it probably would not be profitable to compare the AHN2 with the AHN1 point clouds, because of the difference in resolution. But he would like to be able to compare AHN2 with the future AHN3 data. Pattern analysis is not much sought-after.

5.3.1 Analysis: spatial analyses with other tables

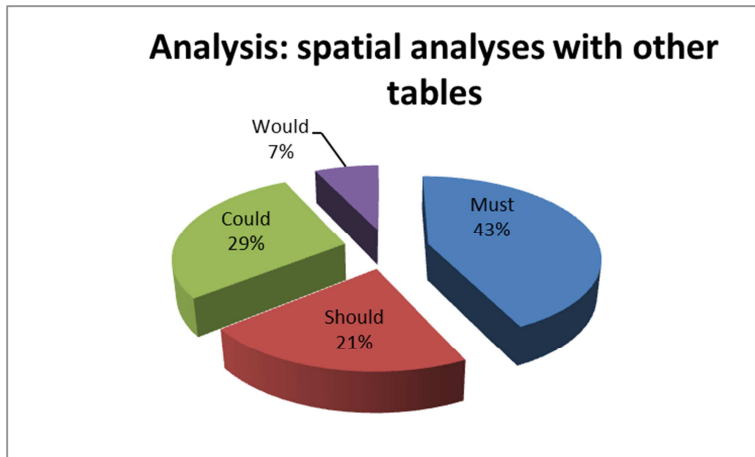


Figure 19 - Spatial analyses with other tables

5.3.2 Analysis: temporal differences

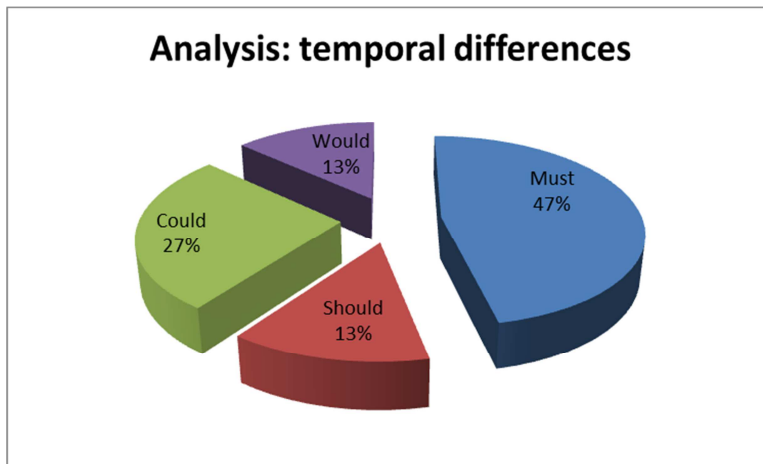


Figure 20 - Temporal differences

5.3.3 Analysis: pattern

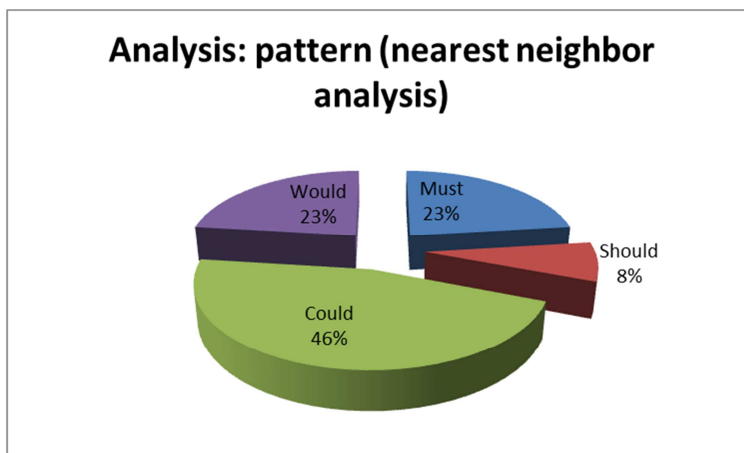


Figure 21 - Pattern analysis

5.4 Access

Most researchers (84 %) want to access the new point clouds spatial DBMS through a web service, either directly through their browser or by letting their desktop software (e.g. ArcGIS) connect to the DBMS. Some researchers - unasked for - mentioned the Web Map Services (WMS) and Web Feature Services (WFS) as protocols to deal with these connections. Freek van Leijen even suggested a configuration (Figure 22). WMS and WFS are not capable of handling raw point cloud data. During the project a new protocol will be developed in order to be able the deal with point cloud data: the Web Point Clouds Services (WPCS) protocol.

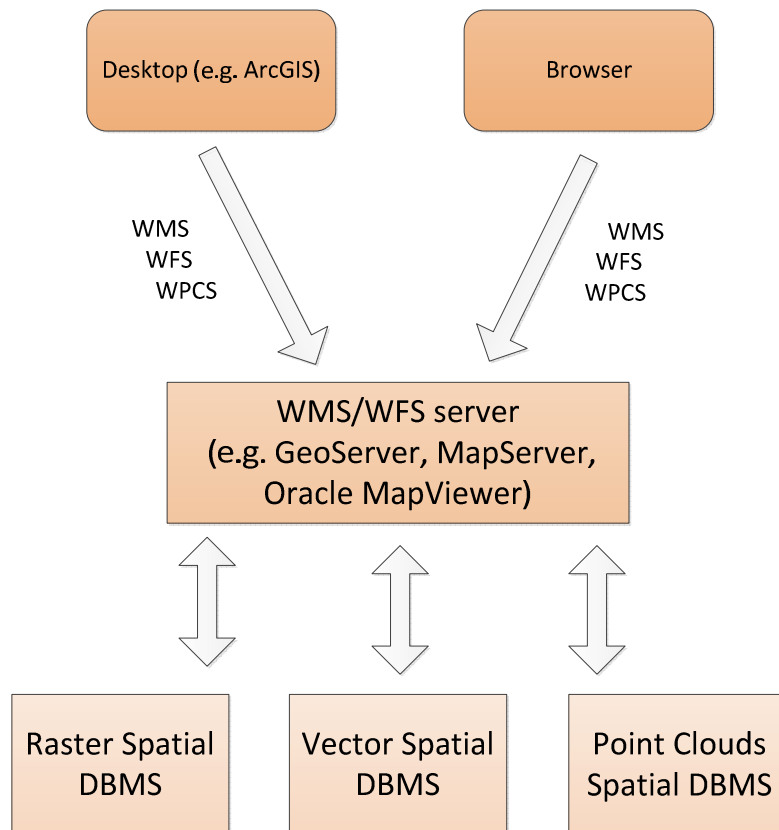


Figure 22 – Schema for access configuration

5.5 View

Like with access, most researchers want to view the selected data in a web application using one of the standard browsers (31%), or by means of a desktop application (37%). Some researchers (14%) also wanted to be able to view the data on a mobile device. One researcher remarked that the viewer should be platform independent.

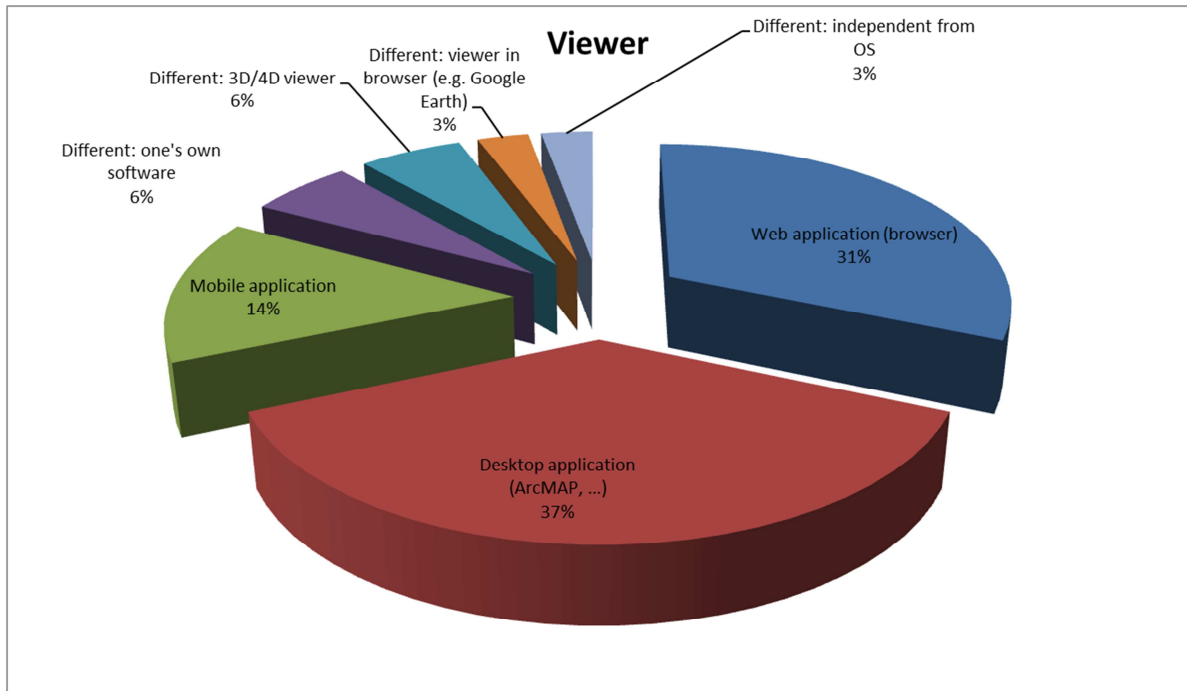


Figure 23 - Preferences for viewing the point clouds

5.6 Navigate

Most researchers want to navigate the point cloud data using traditional techniques like zooming and panning (52%). Some researchers would like to search for an address or a postal code (26%).

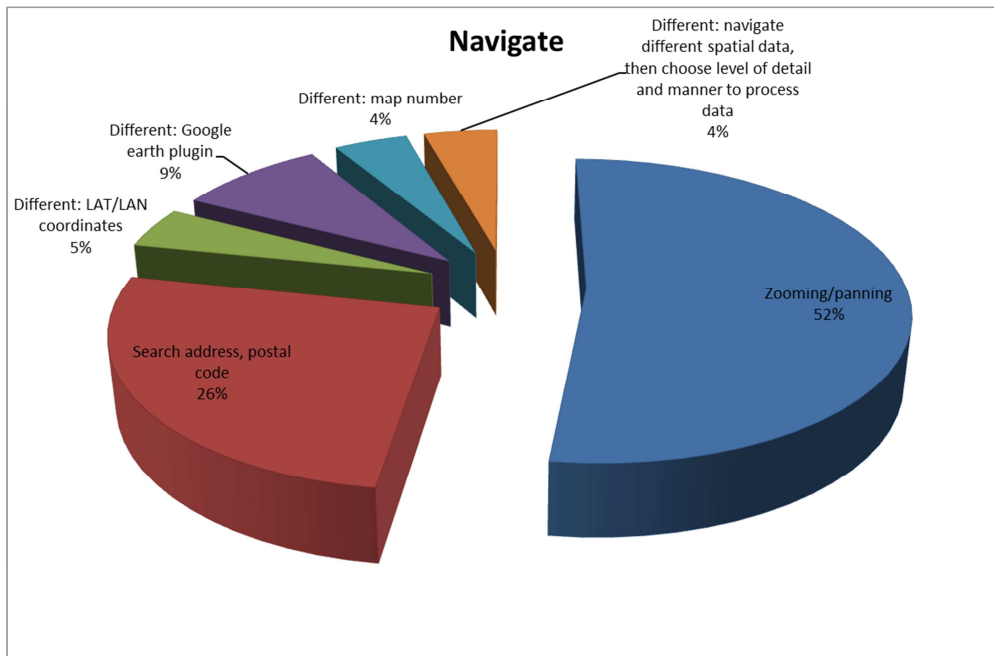


Figure 24 - Preferences for navigating the point clouds

5.7 Download form and format

The raster data is still the most popular format but it is reassuring to find that a lot of researchers also would like to download the point cloud data, if available. A number of download formats have been mentioned. For raster data the researchers would like to use image or publication formats (jpg, pdf) or GeoTIFF and NetCDF format [Website: NetCDF, Website: OpenDAP]. For point clouds downloads in LAS and LAZ formats are the current standard. NetCDF can –in theory- also be used for point cloud data. The new E57 standard has not been mentioned by our respondents [Website: E57].

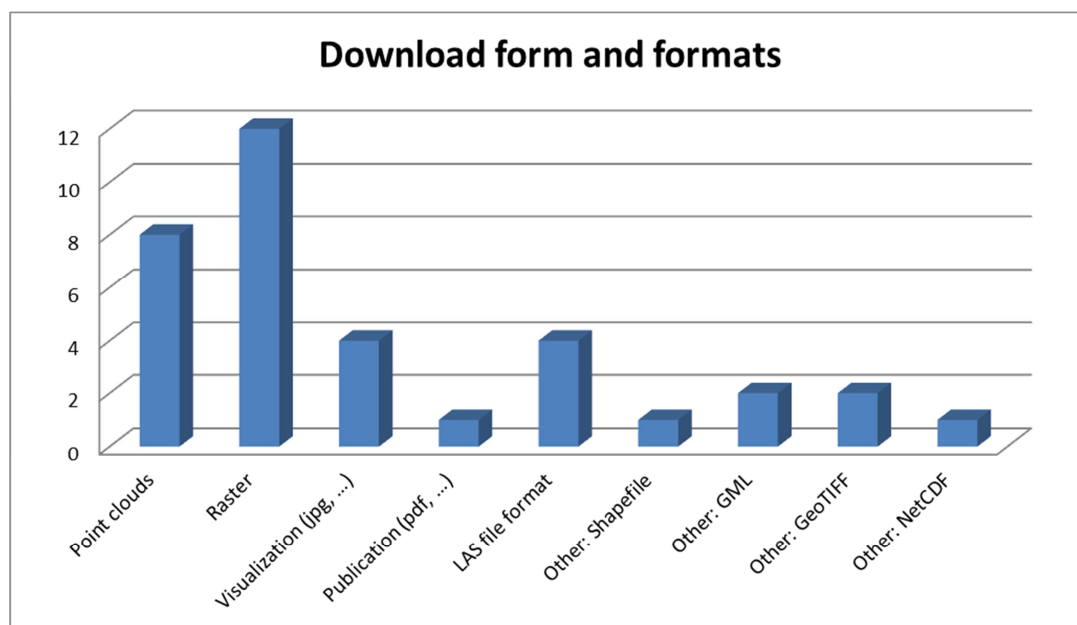


Figure 25 - Data type and download formats

5.8 Attributes

Point clouds can contain additional information that is stored along with every x, y, and z value. The following attributes may be available: RGB (red, green and blue) values, first, second and last pulse recorded, intensity of the pulse, full waveform. Sometimes during post-processing of the data a classification code is added to give a general idea of the object that has reflected the laser pulse [Website: lidar data]. A lot of users would like to use one or more of these attributes in their research, the RGB value being the most popular one.

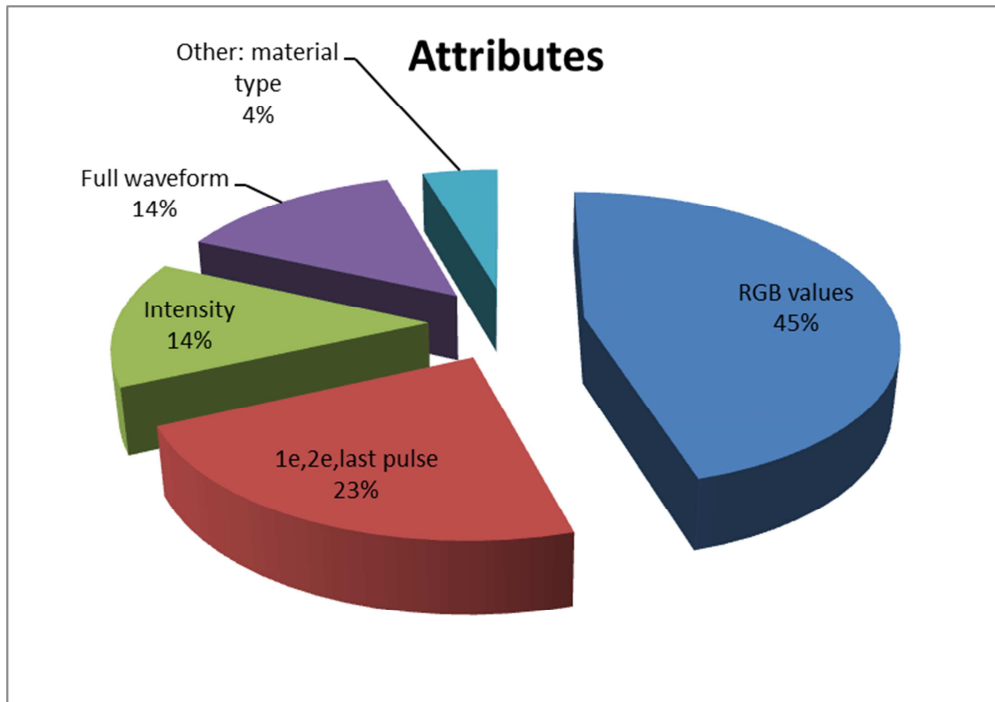


Figure 26 - Preference for type of attributes

5.9 Co-ordinate systems

Most researchers use the Rijksdriehoeksmeting (RD) for locations and the Normaal Amsterdams Peil (NAP) for heights as reference frames. The European Terrestrial Reference System 1989 (ETRS89) is gaining popularity because researchers tend to work more in an international (European) context.

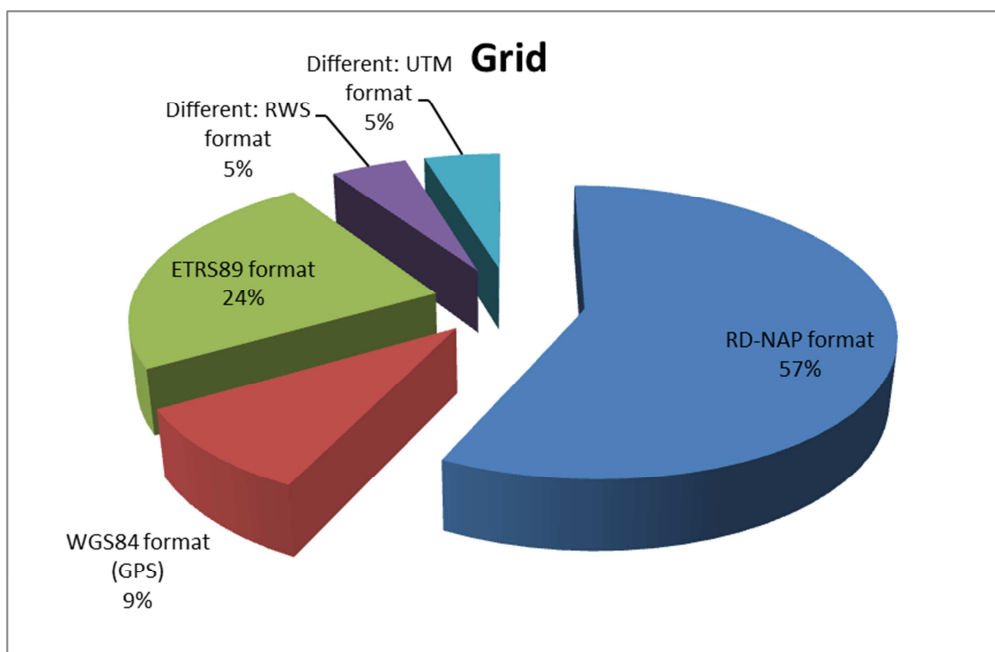


Figure 27 - Preference for reference frames

6 Conclusions

At the moment the AHN raster data are used much more intensively than AHN point cloud data. Most researchers and professionals who were being interviewed in this survey or who completed our questionnaire have used the AHN raster data because (1) AHN raster files are readily available in their organization (2) AHN raster files provide enough detail to carry out their research (3) AHN point clouds are too large to handle. However, most respondents are willing to start working with AHN point clouds if (1) AHN point clouds become more easily available (2) in a format that is easier to handle, because (3) AHN point clouds contain more detailed information and (4) additional information in the form of attributes that are not available in AHN raster data. Some respondents already have successfully used AHN point cloud data in their work.

The respondents would like to access the new point cloud DBMS through a web service, either directly from a browser or from their desktop application (e.g. ArcGIS).

They would like to be able to specify (1) rectangles (2) and polygons to select points from a certain area and (3) perform attribute searches like RGB-values and classification searches like roof tops or dikes. To a lesser degree they want to (4) select point based upon a line and a corresponding buffer zone, e.g. a river or a rail.

They would like to calculate (1) statistical values from a selected area and (2) the (mean) slope orientation and angle of inclination and (3) the point cloud density. To a lesser extent they want to be able to calculate the (4) plane and the (5) point normal estimation. Some respondents want to be able to create profiles and calculate cubic contents of volumes.

Our respondents want to be able to (1) compare point clouds with other data and (2) combine point clouds with older data in order to establish the temporal differences. To a lesser degree they want to do (3) pattern analysis like nearest neighbor analysis.

Once an area or a group of objects is selected, our respondents want to be able to choose between downloading the data in raster format or in point cloud format. If a visualization is made they want to be able to download the visualization as an image or a PDF-file. The respondents will use the downloaded data in their desktop programs (e.g. ArcGIS and GRASS), or with their programming languages (e.g. MATLAB, Python, R) and tools (e.g. sky-view factor tools) to analyze the data.

Researchers at the universities tend to prefer working with open access tools and programming languages, while the professionals from Rijkswaterstaat and Fugro prefer to work with commercial software like ArcGIS and with in-company developed software like Fugro's Lidar processing tools.

Some respondents would like to see some preliminary visualization and analyses done within the point cloud spatial DBMS, in order to get an idea of the data. Once satisfied with the result they will download the data and redo the analysis using their own software. This way, they will be sure that they have absolute control over the results.

7 References

7.1 Publications

P.J.M. Oosterom, W.J.S.M. Wezenbeek (2013). *Massive Point Clouds for eSciences application form*. TU Delft

Santiago Gaitan, Marie-Claire ten Veldhuis, Nick van de Giesen (2013). *Modeling pluvial flooding damage in urban environments: spatial relationships between citizens' complaints and overland catchment areas*. Geophysical Research Abstracts, vol. 15, EGU2013-4851, 2013
<http://meetingorganizer.copernicus.org/EGU2013/EGU2013-4851.pdf>

F.D. van der Hoeven, A. Wandl (2013). *Amsterwarm: Gebiedstypologie warmte-eiland Amsterdam*. Delft Nederland: TU Delft, Faculty of Architecture
<http://repository.tudelft.nl/view/ir/uuid%3Aa93e359e-43b5-416a-9059-2234a7ed89d0/>

Thijs van Lankveld, Marc J. van Kreveld, Remco C. Veltkamp. *Identifying rectangles in laser range data for urban scene reconstruction*. Computers & Graphics 35(3): 719-725 (2011)
<http://www.sciencedirect.com/science/article/pii/S0097849311000434>

H. Ledoux (2009). *Storage and analysis of massive TINs in a DBMS*. TU Delft, GIS technology group.
http://www.ncg.knaw.nl/Publicaties/Groen/pdf/49NCGGroenPointClouds_05.pdf
<http://www.gdmc.nl/events/pointclouds/05%20-%20Hugo%20Ledoux.pdf>

Rutzing, M., Zieher, T., Vetter, M., Geitner, C., Meißl, G., Perzl, F., Markart, G. & Formayer, H. (2013). *Climate induced system status changes at slopes and their impact on shallow landslide susceptibility – a concept*. Alpine Geomorphologie, CH-AT Mountain Days. Mittersill, Austria
<http://www.uibk.ac.at/geographie/lidar/c3s/c3s.html>

C. Steenbergen, W. Reh, S. Nijhuis, M. Pouderoijen (2009). *The Polder Atlas of the Netherlands*. Thoth.

M.W. Straatsma, M.J. Baptist (2008). *Floodplain roughness parameterization using airborne laser scanning and spectral remote sensing*. Remote Sensing of Environment, vol. 112, issue 3, pages 1062-1080.
<http://dx.doi.org/10.1016/j.rse.2007.07.012>

7.2 Websites

E57 file format
<http://www.libe57.org/>

LAStools
<http://www.cs.unc.edu/~isenburg/lastools/>

Lidar data (Esri)
<http://resources.arcgis.com/en/help/main/10.1/index.html#//015w00000041000000>

MoSCoW
http://en.wikipedia.org/wiki/MoSCoW_Method

NetCDF

<http://en.wikipedia.org/wiki/NetCDF>

OPeNDAP

<http://en.wikipedia.org/wiki/OPeNDAP>

SilviLaser conference

<http://silvilaser2012.com/>

Triangulated irregular network (TIN)

http://en.wikipedia.org/wiki/Triangulated_irregular_network

Whitebox Geospatial Analysis Tools

http://en.wikipedia.org/wiki/Whitebox_Geospatial_Analysis_Tools