Quality assessment report to the Danish Elevation Model (DK-DEM)

Andrew Flatman, Brigitte Rosenkranz, Kristian Evers, Peter Bartels, Simon Kokkendorff, Thomas Knudsen, Thorbjørn Nielsen
Introduction

DK-DEM, the Danish Elevation Model, comprises a number of closely related elevation data products with the common aim of helping private and public stakeholders to develop innovative solutions for future challenges, including climate change adaption, urban and rural planning, taxation, and in general any field where information about elevation is useful.

Since 2013, funding for updates of the Danish Elevation Model, DK-DEM, has been made available through the Danish “Basic Data Program” (http://grunddata.dk). The first update (i.e. the second version) of DK-DEM, was published in December 2015, replacing the original 2007 version, which had shown extremely useful for predicting impacts of climate change to the infrastructure, and hence for planning for change in a timely and cost effective manner.

But use of the 2007 version was hampered by a license limiting its use across administrative borders. This was a problem for hydrological modelling, since real water does not respect these boundaries. Hence, enabling a more open access to elevation data was a compelling route to even better hydrological modelling and planning.

The DK-DEM Products

As already mentioned, DK-DEM is more than one product. It is a portfolio of products which are made freely available for both commercial and noncommercial purposes. At the time of writing DK-DEM consists of the following products (see product specifications in Danish at http://kortforsyningen.dk/indhold/data or click on the links below):

- DK-DEM/PointCloud
- DK-DEM/Terrain
- DK-DEM/Surface
- DK-DEM/Rain
- DK-DEM/Searise.

This technical report describes the quality of DK-DEM/PointCloud, the foundation for all derived products.
Quality Control

The scope and objectives of the quality control (QC) process was designed as an integral part of the development of the requirements specification (RS), prior to the public tender for the data. The development of the QC and the RS was guided by two homologous dogma:

- In the RS, do not specify anything we cannot test for
- In the QC, check everything we specify

The QC was further guided by a third dogma, based on the assumption that a large proportion of the checks would be highly automated:

- Avoid spot-checks: Design all checks for full countrywide coverage

This ambitious policy is uncommon, but we believe that its implementation is one of the primary reasons for the very high quality achieved for the new DK-DEM.

Furthermore, the aim of the described QC was that erroneous data would be identified fast and as early as possible in the production flow to increase the final quality of data and decrease time consuming manual processing.

In order to ensure a high and consistent data quality of the second national scale LiDAR mapping of Denmark, the Agency for Data Supply and Efficiency deployed a four layer framework, implementing a continuous data flow from initial point cloud quality screening to final elevation models. The four frameworks are the following:

1. Geometry
2. Classification
3. Finalisation
4. Production

The intention of the QC processes was to strive for the highest possible level of automatisation, like fully automated delivery and geometric control, while the control of the 13 different classes (classification checks) to a higher extent still required manual spot-testing control.
Figure 1) shows different steps in the data flow within the deployed four layer framework including geometry, classification, finalisation and production.
Logistics

The data collection was carried out by airborne LiDAR measurements where up to four airplanes, on cloud free days and within three flying seasons (autumn 2014, spring and autumn 2015) were scanning the Danish landscape. All in all about 415 billion points were collected.

In this project Denmark in its entirety was mapped at four points per square meter in less than two years. This is a large area at a high resolution resulting in lots of flightlines that all needed to be adjusted to fit geometrically. Making all lines fit in one go ensures the best overall geometric match but would at the same time result in an unacceptable long time from flight to delivery. To counter this, the full area was divided into blocks of roughly 50 km x 20 km. The block size was chosen to ensure good geometrical accuracy and being manageable in flight and processing time.

Logistically each block was handled separately, and locked into the national grid using ground control points. This resulted in each block being of similar quality but also that differences accumulate on the borders between the blocks, as can later be seen.

Figure 2) shows the blocks for the project.
Geometry

In the LiDAR project “Geometry” covered all the quality aspects in the raw point cloud data. With this definition the Geometry assessment was divided into the following aspects: Density (distribution), xy-precision, z-precision, xy-accuracy and z-accuracy

Above it can be seen, precision and accuracy were evaluated as two separate aspects of the quality of the geometry. This was due to the fact that LiDAR data, and especially LiDAR data from the Riegl scanners deliver data with very high precision and more typical quality for accuracy. So to have results not skewed by different statistical properties, the data was tested on precision and accuracy separately.

It should be noted that having better precision than accuracy for the raw data is better than having better accuracy than precision. The reason for this is that with good precision, good accuracy can be achieved with good post processing. But having bad precision in the raw data, good precision can never be achieved no matter how good the accuracy is. Hence the specification for this project focussed on equipment delivering very high precision.

Point Density / Point Distribution

Point density and point distribution are two sides of the same issue that together define how well a surface can be modeled, but often distribution is forgotten. In this project it was important that the points were regularly placed on the ground for which reason a scanner using a rotating prism was chosen. This ensured the best point distribution and was as such not directly checked during data-checks.

The point density was designed to be at least 4 points per sq meter and it was specified that this point density was to be achieved for every 100m by 100m in the raw point cloud. The reason for doing the checking in 100x100m squares is to ensure that areas with low point density do not get overlooked in the statistics due to areas with high point density.

The control of point density was done semi automatic in two steps. First all 100x100m squares in a 1x1km tile where automatically checked. If any tiles in a block had any squares below 4 points per square meter, the Tile would be marked for manual evaluation.
Figur 3) shows block C08. The line of red tiles in the western part were delivered in a separate delivery. The residual pink and red tiles are expected outcomes from tiles with water bodies that result in a lower point density.

During the validation phase it has become clear that the main reason for areas not to meet the specification is mainly due to areas covered with water (e.g. wetlands or temporarily submerged fields) or non-reflecting material (e.g. tarmac roofing). All in all the density checks found that the specification was achieved with at least 4 to 5 ppm². Due to overlaps and multiple returns from e.g. vegetation the actual average point density for the entire country is actually above 8 ppm². There was though found one larger area around the city of Aalborg that was lower than the rest of Denmark but still deemed to be just within specification. The reason for this blocks lower point density is that there was a slight layer of moisture on the ground when scanning resulting in less returns from especially horizontal road surfaces and roofs.

Vertical Precision

Often when doing LiDAR strip adjustment and measuring strip offsets, all terrain points in the overlap are used. In this project however, only points on road surfaces were used. This was done out of the idea that bad precision in e.g. vegetation would skew results and using only road surfaces would ensure much better precision and hence much better picture of how well flightlines match. During data control it has become clear that this method of checking is far superior to the regular way of checking strip adjustment as long as there are a sufficient amount of roads in the area.
Specifically for the z precision checks neighboring flight lines were automatically interpolated onto each other on flat road centers and statistically compared. The results were then plotted per block and manually evaluated.

Figure 4) shows the result from block C08 comparing neighboring flight lines on road surfaces in the overlap area.
Horisontal Precision

As for z precision checking the xy precision was in this project done on specially selected objects only. In this case gabled building roofs. Gabled roofs were chosen due to the nature of LiDAR and the fact that LiDAR point clouds are discrete points making it hard to compare directly. For more information please see appendix.

For the xy precision check the roof ridge was first automatically created from the point cloud. These roof ridges were then compared automatically to matching roofridges from the neighbouring flightlines. As for z precision the results from the xy precision checks were then plotted per block and manually evaluated.

*Figure 5* shows the output of the QC of first and second delivery of block A2A2. The red dots in the plot on the left side (first delivery) shows buildings where the ridges are displaced more than 15 cm. This delivery was rejected and the supplier carried out another strip adjustment, which results in the plot on the right hand side. This delivery was accepted. The corrected point cloud still have some
Vertical Accuracy

For checking the Z accuracy, Ground Control Points (GCP) are placed on plane horizontal surfaces. Often these would be middle of carparks, driveways or equal. Often in LiDAR projects one GCP is actually 16 - 36 points measured in a mesh. In this project one GCP was in fact only one point. In a sparsely populated point cloud it is necessary to have many GCP’s in a single GPS mesh for ensuring GCP’s close to the LiDAR points. This was not the case in this dense point cloud with a distance of less than 40 cm between the LiDAR points. So instead of having to measure one GCP mesh many times, more single GCP’s around the country were measured and hereby given a much better foundation for
In each block the Z accuracy was calculated against 40-70 GCPs.

Count:75
Minimum value:-0.065064
Maximum value:0.062115
Mean value:0.001675
Median value:0.002683
Standard deviation:0.033022

*Figure 6*) shows the result of the Ground Control Point (GCP) statistic of block C10.
Horizontal Accuracy

For checking the XY Accuracy roof ridges were again used. The roof ridges form the point cloud were calculated as when checking xy precision. These roof ridges were built from the point cloud and compared to GeoDanmark buildings. These points were transformed onto the photogrammetrically measured polygon. The result of the transformation were reported. GCP were surveyed and compared to the point cloud. Road centerlines were interpolated from the point cloud and compared to GeoDanmark data.
Count:18480
Minimum value:0.000108
Maximum value:0.699865
Mean value:0.173394
Median value:0.115581
Standard deviation:0.166632

Figure 7) shows the result of the roof ridges calculated from the point cloud compared to GeoDanmark buildings.

Block Edges (zipper)
As mentioned earlier when discussing logistics dividing the country into blocks often result in discrepancies cropping up on the borders of the blocks and needs to be handled specifically. To handle this the last strip of tiles in each block was grouped as a special Zipper-block that could be checked in the same way as the regular blocks going through all the same checks.

Count:169816
Minimum value:0.000374
Maximum value:0.832864
Mean value:0.035125
Median value:0.028085
Standard deviation: 0.023488

*Figure 8*) shows the vertical precision results of the zipper-block that was handled as the last block.

As was expected it was found that the errors were larger in the Zipper region than in the regular blocks. This is of course problematic but not much could be done within the time frame of the main project. But what has been learnt during the process is that especially the borders between either blocks in a single project or between new data and old data needs to be handled as early as the flight planning stage. If the planning of the LiDAR flight mission does not take borders into account it becomes almost impossible to fix later.

**Classification**

The data was running through an automatic test with software developed for this purpose. Unlikely situations like vegetation inside buildings, buildings on top of roads and others were detected. Human experts interpreted the results by using a stepping function that was developed for QGIS. These humans in the loop were stepping through all detected objects. The main focus was kept on the quality of the products DK-DEM/Terrain and DK-DEM/Surface, as they are seen as the end products. Keeping in mind that the classification was done automatically, we were accepting classification errors that would not affect the two raster models, for instance “vegetation” points that are classified as “building”.
While common quality control processes for classification of laser scanned data are based on manual tasks where spot-check samples have been examined by an operator comparing with supporting data (eg. ortho imagery, topographic maps etc), the classification control for the DK-DEM was to a large degree based on nation wide tests with statistical results. Based on these outcomes decisions for acceptance, rejection or further investigation were taken.

The point cloud includes the following classes:

<table>
<thead>
<tr>
<th>Class</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>Created, never classified</td>
</tr>
<tr>
<td>1</td>
<td>Surface. Processed but unclassified</td>
</tr>
<tr>
<td>2</td>
<td>Terrain. Bare earth ground</td>
</tr>
<tr>
<td>3</td>
<td>Low vegetation. 0 - 0,3 m</td>
</tr>
<tr>
<td>4</td>
<td>Medium vegetation. 0,3 - 2 m</td>
</tr>
<tr>
<td>5</td>
<td>High vegetation. &gt;2 m</td>
</tr>
<tr>
<td>6</td>
<td>Buildings, structures. Buildings, houses, silos etc</td>
</tr>
<tr>
<td>7</td>
<td>Outliers. Spurious high/lown point returns and noise (unusable)</td>
</tr>
<tr>
<td>8</td>
<td>Model key points</td>
</tr>
<tr>
<td>9</td>
<td>Water. Surface water</td>
</tr>
<tr>
<td>10</td>
<td>Ignored points (breakline proximity)</td>
</tr>
<tr>
<td>14</td>
<td>Wire – Conductor (Phase)</td>
</tr>
<tr>
<td>17</td>
<td>Bridge deck.</td>
</tr>
<tr>
<td>18</td>
<td>High noise</td>
</tr>
<tr>
<td>19</td>
<td>Terrain in buildings</td>
</tr>
<tr>
<td>20</td>
<td>Low vegetation in buildings</td>
</tr>
<tr>
<td>32</td>
<td>Objects manually excluded from surface points</td>
</tr>
</tbody>
</table>

Some notes to the classification.
- Class 5 (High vegetation) was used as a “catch all” class when no other class could automatically be assigned. For instance some tall structures not detected as “building” (towers, antennae etc) ended up in class 5.
- Class 9 (Water) is difficult to get right. The amount of points for a given water body is determined by many parameters (e.g. wind, scan angle, position relative to the strip, algae, ...) making it difficult to ensure right coding.
- Class 10 and 32 are not used in the final delivery.
Furthermore, the following information is stored for each point:
- Elevation
- Class
- Intensity
- Scanline ID
- Position in the reflection signal (first, ..., last reflection)
- RGB-value, for the points collected by daylight.

Automatic Classification Control

For this project the following automated tests were carried out.

Point Classification Classes

A first basic check of the point cloud classification was counting the amount of each class that was present in each 1km by 1km tile. By visualising the data as shown in the figure below, it was possible detect suspicious pattern of classification in the delivered data.

*Figure 9*) shows block A1B1 with each tile colored according to the amount of points within the named class. Red tile mean that there were no points, shades of blue indicate increasing amount of points of the named class. In the bottom right figure the western part of the block indicates that no bridge points were present. Visual inspection proved this to be correct.
Comparison between the new and the 2007 Terrain Model

A comparison with the 2007 terrain model showed differences in the two datasets. As it can be seen from the figure below some general differences revealed between the datasets. Often this is caused by inaccurate strip matching in the DK-DEM-2007 dataset. Also some agricultural fields were highlighted revealing classification errors (top of crops are classified as terrain instead of vegetation). Most of the differences were pointing to erroneous classification in the 2007 dataset but some issues were also found and corrected in the new data.

Figure 10) shows a plot of differences between the 2007 and the new version of DK-DEM. Dark red and dark blue colours represent differences.
Spike Check

The spike check was used to identify suspicious outliers in the point cloud. The criteria for a point to be classified as a spike was that the interpolation between points (search radius 1.5m) has an angle of 25 degrees towards all neighbouring points. Less than 2 spikes/km² were generally accepted for a standard block.

![Spike Check Image](image1.png)

*Figure 11) shows all identified spikes in block A1B1.*

If the statistical pattern was suspicious the reason was investigated and erroneous points were reclassified from “terrain” to “unclassified”. The reclassification process is discussed in the chapter Finalisation.

![Spike Examples Image](image2.png)

*Figure 12) shows examples of detected spikes on jetties and ships masts.*
Vegetation below Building Polygons

GeoDanmark buildings polygons that contained points classified as vegetation were pointed out and evaluated. Vegetation points inside building polygons were reclassified to outliers in the finalisation process (see chapter Finalisation).

*Figure 13*) shows an example where the number of vegetation points inside a building polygon exceeds the given threshold. Red buildings are above threshold and green below.

Building Points in Polygons

GeoDanmark building polygons that did not contain points classified as buildings were pointed out and evaluated. Buildings, with an area larger than 1600 m², that were classified wrong were re-classified manually. In many cases it turned out that the buildings were demolished and the GeoDanmark data were outdated, as seen in the figure below.

*Figure 14*) shows an example of a building that was demolished and where the GeoDanmark data was outdated. It was pointed out correctly the test.
Auto Building Check

Errors in the classification resulted in tree canopies classified as buildings. By polygonising the point class “building” a representation comparable to the GeoDanmark data building polygons was derived. The two datasets were compared and the result evaluated.

Figure 15) shows auto-generated building polygons derived from the point cloud. The data was accepted after a comparison with GeoDanmark building polygons.

Manual Classification Control

The results from the automatic classification control were evaluated by human experts. When the automatic control had been accepted a visual inspection of the data was carried out. The visual inspection was based on hillshaded DEMs and classification grids based on the point cloud classification. An example is shown in the figure below where each cell of the hillshade is coloured according to the predominant classification in that cell.

Figure 16) shows the result of a wrong classification.  
Left: Dikes and ditches are classified as building (red area) and are resulting in an incorrect terrain model (holes in terrain).  
Right: Corrected terrain model.

Trained experts were using the QGis stepping function and controlled visually every 1km x 1km hillshaded tile. The errors found here were corrected in the finalisation step, either by automatic or by manually corrections.
Finalisation

In the finalisation step corrections to the point cloud were applied. The corrections utilised both automatic and manual procedures to correct errors in the point cloud. After each of the correction processes the corrected point cloud tiles were stored in the data management system in order to keep track of changes in the dataset.

Automatic Corrections

The automatic corrections, or patching procedure, took a number of inputs from the classification control and corrected them as good as possible. Furthermore the patching procedure filled relevant holes void of data in the point cloud with data from the previous version of DK-DEM-2007.

Filled in Data Voids

The delivered point cloud exhibited a number of gaps in the data coverage. This was due to blank rainwater on soaked fields. The holes in the point cloud were detected by looking for area with low point density and using a data void (floodfill-like) algorithm to find the full extent of the area with no data. Low or no data coverage under buildings and other large features that covers the terrain were ignored when searching for data voids. The data voids that were found were filled with data from the 2007 point cloud. This way we ensured a consistent model across the whole country.

Figure 17) shows the result of the data void fill algorithm.  
Left: Detected area with no data coverage (orange area).  
Center: Terrain model based on the unmodified data. Notice the big triangles in the middle, which are a result of the missing data.  
Right: Terrain model after data from the old point cloud has been added to the new model.
Reclassify Spikes

Spikes identified in the classification control were corrected in the finalisation step. All identified spikes that differed more than 0.75 m from the surrounding terrain were reclassified as outliers. Only points classified as terrain were considered when detecting spikes.

Figure 18) shows the effects in a terrain model before and after spikes were reclassified. Left: Before the outliers were reclassified. Right: After the outliers were reclassified.

Reclassify floating Objects

Floating objects, such as high tension wires and bird flocks, are for obvious reasons unwanted in the terrain and surface model. Floating objects were in many cases classified by the supplier as vegetation in the data as these objects share some of the characteristics that are used in the automated classification procedure. By using a voxel filtering method we were able to reclassify most of the floating objects in the point cloud.

Figure 19) shows examples of surface model before and after the floating-objects-filter has been applied on high tension wires can be seen. Left: Before the voxel filter has been applied. Right: After the voxel filter has been applied
In most cases floating objects were reclassified as “high noise”. Supporting vector information was imported to the algorithm to avoid erroneous reclassification of objects that were not considered noise. Due to alike characteristics of floating objects and trees, points in forests and close to single trees were not taken into account. Known high tension lines were classified as conducting wires (class 14).

Figure 20) shows bird flocks in the surface model.  
Left: Before the voxel filter has been applied. Right: After the voxel filter has been applied.

Reclassify Terrain and Vegetation Points inside Buildings

Due to the nature of the classification algorithms used by the data supplier, in some cases parts of buildings were classified as vegetation or terrain. Buildings classified as vegetation are obviously wrong, but it doesn't affect the derived terrain and surface models. Buildings classified as terrain, on the other hand, affects the terrain model as seen in the left part of the figure below. In this case the terrain points are below ground level because of multipath return signals, and should have been classified as “low noise”.

Figure 21) shows buildings classified as terrain.  
Left: A large triangle is generated because multipath return points are classified as terrain.  
Right: The points are reclassified to low noise and excluded when generating the terrain model.
The wrongly classified points were identified by looking for points classified as terrain and low vegetation inside GeoDanmark building polygons. All terrain and low vegetation points found inside the polygons were reclassified as “terrain in building (19)” or “low vegetation in building (20)”.

Manual Corrections

Manual editing of the point cloud were used to fix the most significant classification errors not handled by the automatic corrections. Manual classification corrections had been applied where the manual classification control found significant classification errors. At the time of writing roughly two percent of the tiles in the point cloud had been edited manually.
Conclusion

Updating and improving the DK-DEM was a procurement that took more than 3 years. The first step in the procurement process was to collect customer needs and translate them to the product specification. The next steps were choosing the supplier, planning and carrying out the data acquisition. The data acquisition had been carried out within 3 acquisition seasons by up to 4 airplanes scanning the Danish landscape. Even though these steps required very different competences, all of them were equally important for accomplishing the successful data procurement and archiving the good data quality that satisfies the user's needs.

Within this project we made two untraditional strategic decisions which showed out to be very successful.

The first untraditional strategic decision that was taken early in the process was that the we did not procure the the actual grid models, DK-DEM/Terrain and DK-DEM/Surface, but produced these ourselves.

This was due to the necessity of having full control over the interpolation method when producing either derived products or updating the DK-DEM. This decision opened up for the development of a production system where products, like e.g. the hydrological adapted elevation models, can be derived and where DK-DEM and related products efficiently can be updated. This makes us independent from the supplier of the elevation model and gives us the possibility to update DK-DEM whenever we detect changes of the terrain that make this necessary.

The second untraditional strategic decision that was taken in the assessment process was that we would

1. **not specify anything that could not be tested**

and furthermore

2. **control everything that was specified as a nation wide control and not only as a spot-check**

This practice was - at the time of writing - quite unusual in the geodata business. Nevertheless, it turned out to be very successful for two reasons. The requirements were clear and easy to understand by the supplier and we were sure that we received what we had ordered, not only at spots, but for the whole country. It is our conviction that the achieved high data quality of DK-DEM is resulting from implementing this strategy.

Both parties, the Agency for Data Supply and Efficiency and the supplier Aerodata Netherlands, worked in a close cooperation on continuously improving the production methods, having in mind to reach the highest possible data quality. For that reason the Agency of Data Supply and Efficiency was improving the classification of the data even more than what was delivered by the supplier, by taking advantage of local knowledge.
The data complies with high quality requirements that we put in prospect in the procurement phase. While most areas have an even better quality than ordered some areas, often block edges, have a somewhat lower quality. As this was expected (errors are often accumulating at edges) and as the overall statistic is within specified quality these areas were accepted.

We firmly believe that this high quality data will enable users to utilize data in new, different and more advanced and innovative ways. Having the potential of and need for data in mind, our goal is to continue developing methods and data and to acquire data at a level that DK-DEM always will be as up to date and usable as possible.