3.4 Classification

TerraSolid's TerraScan software was used to classify the raw laser point into the following categories: ground, non-ground (default), aerial points and low points.
Because of the large size of the LiDAR data the processing had to be done in tiles. Each survey segment was imported into TerraScan projects consisting of 1000m x 1000m tiles aligned with the 1000 units in UTM coordinates.

The classification process was executed by a TerraScan macro that was run on each individual tile data and the neighboring points within a 40m buffer. The overlap in processing ensures that the filtering routine generate consistent results across the tile boundaries.

The classification macros consist of following algorithms:
1) *Isolated points*: This routine classifies points which do not have very many other points within a 3D search radius. This routine is useful for finding isolated points up in the air (fog) or below the ground (multipath). When possibly classifying one point, this routine will find how many neighbouring points there are within a gived 3D search radius. It will classify the point if it does not have enough neighbours.

2) *Air points***:** It classifies points which are clearly higher than the median elevation of surrounding points. It can be used to classify noise up in the air. When possibly classifying one point, this routine will find all the neighboring source points within a given search radius. It will compute the median elevation of the points and the standard deviation of the elevations. The point will be classified only if it is more than a certain limit (user defined) times the standard deviation above the median elevation. Comparison using standard deviation results in the routine being less likely to classify points in places where there is greater elevation variation.

3) *Removal of "Low Points"*. This routine was used to search for possible error points which are clearly below the ground surface. The elevation of each point (=center) is compared with every other point within a given neighborhood and if the center point is clearly lower then any other point it will be classified as a "low point". This routine can also search for groups of low points where the whole group is lower than other points in the vicinity. Input parameters used were:

4) *Ground Classification*. This routine classifies ground points by iteratively building a triangulated surface model. The algorithm starts by selecting some local low points assumed as sure hits on the ground, within a specified windows size. This makes the algorithm particularly sensitive to low outliers in the initial dataset, hence the requirement of removing as many erroneous low points as possible in the first step.

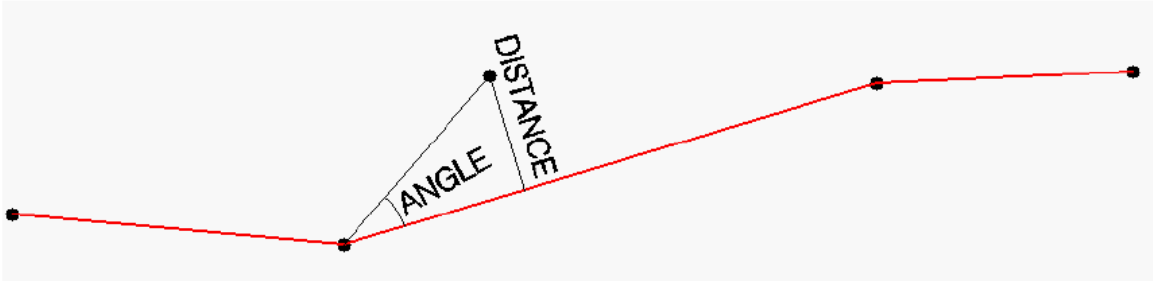

Figure 6 Ground classification parameters

The routine builds an initial model from selected low points. Triangles in this initial model are mostly below the ground with only the vertices touching ground. The routine then starts molding the model upwards by iteratively adding new laser points to it. Each added point makes the model follow ground surface more closely.

The various input parameters are

Max Building Size (window size): 40.0m
Max Terrain Angle: 88.00 deg.
Iteration Angle: 6.0 deg
Iteration Distance: 1.4m

These parameters depend on the properties of the area such as extent of urbanization, vegetation density and terrain (flat/rugged). Iteration parameters determine how close a point must be to a triangle plane so that the point can be accepted to the model. Iteration angle is the maximum angle between point, its projection on triangle plane and closest triangle vertex. The smaller the Iteration angle, the less eager the routine is to follow changes in the point cloud. Iteration distance parameter makes sure that the iteration does not make big jumps upwards when triangles are large. This helps to keep low buildings out of the model.

3.5 DEM production

The point data is output from TerraScan in 1000m x 1000m tiles, with 40m overlap. Two sets of files are generated, in XYZ ASCII format: filtered (ground class) and unfiltered (ground and “default” classes). In the unfiltered dataset the outlier classes are excluded from output (aerial and low points). The overlap is needed in order to generate a consistent interpolation across tile edges and it will be trimmed in the final tile DEMs.

A set of tiles in the “comprehensive” format is also outputted, to be used by the GEON online distribution and processing center. The various file formats and file naming conversions are described in the next section.

The point tiles are gridded using Golden Software’s Surfer 8 Krigging at 0.5m cell size, using a 5m search radius for the unfiltered point data and 25m for the filtered.

The griding parameters are:

Gridding Algorithm: Kriging
Variogram: Linear

Nugget Variance: 0.15 m
MicroVariance: 0.00 m
SearchDataPerSector: 7
SearchMinData: 5
SearchMaxEmpty: 1
SearchRadius: 5m (unfiltered), 10m (filtered)

The resulting tiled Surfer grid sets are transformed using in-house Perl and AML scripts into ArcInfo binary seamless tiles at 0.5m cell size. Due to the large area covered by the segments and the ArcInfo software limitations it is not possible to create one large mosaic for the entire area so the 0.5m tiles are mosaiced at 1m resolution into 10Km wide segments.

The point tiles are the corresponding grids and mosaics are all positioned in the ITRF2000 reference frame and projected into UTM coordinates Zone 10N, all units in meters. The elevations are heights above the ellipsoid. Because ArcInfo doesn't support directly this particular projection, the grids are assigned the following projection information: UTM Zone 10N, WGS84 (original) datum.