TESTING THE IMPACT OF 2D GENERALISATION ON 3D MODELS – EXPLORING ANALYSIS OPTIONS WITH AN OFF-THE-SHELF SOFTWARE PACKAGE

E. Muñumer Herrero[1], C. Ellul[1], J. Morley[2]

[1] Department of Civil, Environmental and Geomatic Engineering, University College London, UK
(estibaliz.herrero.17, c.ellul)@ucl.ac.uk
[2] Ordnance Survey, Southampton, UK – Jeremy.Morley@os.uk

KEY WORDS: 3D generalisation, 3D models, 3D buildings, simplification, aggregation, performance

ABSTRACT:

Popularity and diverse use of 3D city models has increased exponentially in the past few years, providing a more realistic impression and understanding of cities. Often, 3D city models are created by elevating the buildings from a detailed 2D topographic base map and subsequently used in studies such as solar panel allocation, infrastructure remodelling, antenna installations or even tourist guide applications. However, the large amount of resulting data slows down rendering and visualisation of the 3D models, and can also impact the performance of any analysis. Generalisation enables a reduction in the amount of data - however the addition of the third dimension makes this process more complex, and the loss of detail resulting from the process will inevitably have an impact on the result of any subsequent analysis.

While a few 3D generalization algorithms do exist in a research context, these are not available commercially. However, GIS users can create the generalised 3D models by simplifying and aggregating the 2D dataset first and then extruding it to the third dimension. This approach offers a rapid generalization process to create a dataset to underpin the impact of using generalised data for analysis. Specifically, in this study, the line of sight from a tall building and the sun shadow that it creates are calculated and compared, in both original and generalised datasets. The results obtained after the generalisation process are significant: both the number of polygons and the number of nodes are minimized by around 83% and the volume of 3D buildings is reduced by 14.87 %. As expected, the spatial analyses processing times are also reduced. The study demonstrates the impact of generalisation on analytical results – which is particularly relevant in situations where detailed data is not available and will help to guide the development of future 3D generalisation algorithms. It also highlights some issues with the overall maturity of 3D analysis tools, which could be one factor limiting uptake of 3D GIS.

1. INTRODUCTION

Popularity of three-dimensional (3D) city models has increased massively in recent decades. While the early uses of 3D city models were mainly dominated by visualisation (Christen and Nebiker, 2015; Glander and Döllner, 2009), e.g. 3D urban visualisations for disaster management (Kemec et al., 2010), as the technology developed, they have been useful for many purposes beyond visualisation (Biljecki et al., 2015), for example solar radiation distribution calculations (Hofierka and Zlocha, 2012), noise impact studies (Stoter et al., 2008), 3D Cadastre (Shojaei et al., 2017; Stoter and Ploeger, 2003), urban infrastructure planning (Herbert and Chen, 2015) and so on. 3D City Models - showing details of buildings (indoors and outdoors), roads, parks, street furniture – are also fundamental to enabling Smart Cities. They link information such as traffic flows, pollution, tourism, utilities, infrastructure and public transport to address real world problems in a wide variety of disciplines.

While in theory the same 3D model could be used for these multiple applications, in reality for some applications a high level of detail is required whereas for others this is not the case. Different details are also required for different applications - a tourism application might focus on details about landmarks, whereas a solar panel application might require detailed roof structures. In order to make 3D city modelling efficient and reusable to address different user needs, all the information should be derived from one detailed source – this is known as generalisation.

A number of attempts have been made in the past to develop algorithms for the automatic generalisation of 3D buildings (see Section 2.2), However transitioning from a detailed three-dimensional representation to a coarser one maintaining both geometric and semantic characteristics is still challenging, difficult and time consuming. The process of developing such an algorithm would be assisted by having greater clarity about the end goals – i.e. what the resulting generalised dataset should look like. As part of this, it is important to understand the impact of using generalised data with 3D applications.

To further this understanding, this paper presents an introductory study of the impact of the generalisation on the results of two algorithms – line of sight and shadow casting from the buildings. In the absence of commercially available 3D generalisation algorithms, traditional 2D cartographic generalisation operators are used, in combination with extrusion to answer the following question:

What is the impact of generalisation on the results of 3D line-of-sight and shadow-casting algorithms? We consider this question from the point of view of the variation in analytical results obtained as well as by examining the impact on algorithm performance, and the results will:
- Provide insight into the impacts of using a generalised dataset for these operations, which is particularly relevant for situations where more detailed data is not available.
- Guide further work into 3D generalisation by providing a preliminary insight into the consequences of loss of detail

on the algorithms tested. This in turn will inform decisions as to which elements of 3D data can be generalised without consequence and which are important to retain.

- Additionally, as we will be using off-the-shelf tools for this task we will gain initial understanding related to the maturity of 3D algorithms.

2. LITERATURE REVIEW

2.1 Complexity of 3D Data

Several techniques and methodologies are available to create 3D models including processing data from photogrammetry, remote sensing or laser scanning (Lafarge and Mallet, 2012). When the requirements for detail are not very high, elevating the 2D footprint data to a given height, known as extrusion, is an efficient process to model buildings of a city in 3D (Ledoux and Meijers, 2011). However, even using this simple approach, 3D models can be complex – e.g. Wong and Ellul (2016) note that an LOD1 model of Toronto contains 397,602 buildings and over 10 million vertices (covering 709 km^2) and the 3D model of the city of Berlin (LOD2) 537,208 buildings and over 10.5 million vertices (for 890 km^2). They also note, the method of capture is also highly correlated to the overall geometric complexity and size of the model (ibid).

2.2 Generalisation

Generalisation is a process that categorises features and excludes unnecessary detail in order to reduce visual complexity (Robinson et al., 1995). In traditional cartography, reducing reality to a given map scale while the most important features are emphasised makes generalisation necessary: distances, lengths and widths are shortened and adjacent objects get merged.

Generalisation can be said to have multiple purposes – firstly, to ensure that map information is presented in an understandable way for a map user - i.e. that the correct, relevant details are shown but that the map is not overcrowded as the scale changes; secondly – to create a dataset that is suitable for various analytical tasks – balancing between the computational complexity required to analyse very detailed datasets, and the loss of information (and hence accuracy of results) if the data is over-generalised (ibid). In all cases, generalisation also permits a 'create once, use many times' approach to data – which maximises the return on investment in data capture.

2.2.1 Map Generalisation Steps
The first operation to be performed is the selection of necessary features and attributes, which will depend heavily on the purpose for which the new map is being created. Once this process is completed, drawing the objects at the given map scale is required –this is carried out by generalisation (Robinson et al., 1995) using a combination of the following steps:
- Classification: ordering and grouping features by their type.
- Aggregation: substituting multiple features into a single one.
- Simplification: eliminating unwanted detail.
- Exaggeration: enhancing the important characteristics.
- Symbolization: replacing features with symbols.
- Induction: deduction of the relationships among features.

Generalisation must preserve a harmony and balance between the retained and omitted data, always dependant to the level of scale of the outcome (ibid).

2.3 3D Generalisation

2.3.1 Conceptual Approaches – CityGML and LoD
Many three-dimensional city models are represented and exchanged in the open data model CityGML, based on the international standard for spatial data agreed by the Open Geospatial Consortium (OGC) (Gröger et al., 2012). The standard permits City Model representation not only through graphical means but also in relation to semantic characteristics, which are also considered for thematic applications. City models are represented using five different levels of detail (LoD0-LoD4). Buildings and building parts are represented from LoD1 - a block model with a flat roof to - LoD4 - a building with detailed façades and interiors with windows and doors. Generalisation operators are important when converting the model from certain LoD to a lower one.

2.3.2 Previous Implementations of 3D Generalisation
While 2D generalisation algorithms are available in commercial software, 3D generalisation is still in its infancy. Much of the implementation work on 3D generalisation algorithms has taken place within an academic environment, with relatively little work taking place recently. Implementations vary as to whether they take into account the different components of a 3D model – geometry, semantics and texture, and whether their outcome is aimed towards visualisation or analysis. Table 1 summarises some of the key approaches, with further detail of each in the subsequent paragraphs, with some key approaches summarised below.

Approaches (earliest first)	Geometry	Semantics	Textures	Storage	Visualization	Analysis
Kada (2002)	✓	✓	✓	✓	✓	✗
Sester and Brenner (2005)	✓	✗	✗	✗	✓	✗
Forberg (2007)	✓	✗	✗	✗	✓	✗
Fan et al. (2009)	✓	✓	✓✗	✓	✓	✓
Fan and Meng (2012)	✓	✓	✓✗	✓	✓	✓
Baig and Rahman (2013)	✓	✓	✗	✗	✓	✗

Table 1. Objectives of existing approaches.

Kada (2002) developed an algorithm that first creates a constraint building model in where the faces of the model are grouped by coplanarity, parallelism and rectangularity constraints hierarchically. Afterwards, features such as extrusions are detected and their significance over the global look of the model is evaluated. The features of least significance are then eliminated and the constraint building model altered. Finally, a new location for the vertices of the constraint building model is calculated by least squares adjustment. In many cases overall complexity fell by between 30% and 50% (ibid).

In order to obtain a continuous generalisation for use in a mobile device with a small display, Sester and Brenner (2005) proposed elementary generalisation operations (EGO's) as the key tools for the generalisation of building ground plans and a typification of buildings. The sets of EGO's presented are applicable from a detailed level to a coarser one and also applicable in the inverse. Further reduction of detail is acquired with the operation of the amalgamation and operations to remove offsets, extrusions and corners. This is followed by a discrete process known as typification, which replaces objects with more 'typical' representations. Using this approach an object is gradually

modified rather than being replaced as it moves from one level of detail to another.

An experiment on 3D generalisation based on a scale-space theory from image analysis is performed by Forberg (2007). Different representations at different scales are derived from an image by a scale-space and thus this theory is suitable to generate different levels of detail for 3D models. Forberg (2007) presents results relating to simplification of orthogonal structures and suggests squaring as appropriate for roofs and walls having non-orthogonal structures.

The algorithm presented by Fan et al. (2009) extends Sester and Brenner (2005), taking the semantic information into account in order to avoid deleting important features and the aggregation of polygons which belong to different objects, as well as applying a typification process. As a result, a notable 90% reduction of the storage space was obtained substantially from the original 3D model to the extracted shell, without losing the overall appearance and semantic information of the buildings.

Fan et al (2009) do not include edges and vertices in their process. This omission is addressed by Fan and Meng (2012) who describe a three-step process which starts with the exterior shell of the 3D building extracted using methods proposed by Fan et al., (2009). Roof structures are generalised separately and they take distortions of edges and angles of roof polygons when projected onto the ground into account, by simplifying these on a plane rather than at ground level. However, this has not yet been developed to allow the algorithm handle further extrusions beyond the initial extrusion.

More recently, Baig and Rahman (2013) extend the work of Sester and Brenner (2004) and Fan et al. (2009). Semantic information, height and positional accuracy of the 3D objects are considered in order to derive multiple LoDs and the simplification methods used -removal of intrusions, extrusions, offset and corners and aggregation of footprints- are based on neighbouring edges with semantic rules are imposed to avoid deleting important objects. The authors highlight the semantic-based removal of building parts is the strength of their algorithm for applications in where maintenance of important parts is the main priority.

2.3.3 Testing Performance Impact of 3D Generalisation

As noted above, one of the main outcomes of a generalisation process is to reduce the quantity of data required for analytical tasks. Ellul and Joubran (2012) use 2D generalisation operators - simplification and aggregation - in order to create a generalised 3D city model by extrusion of the dataset, focussing on testing the impact of generalisation on rendering performance. The results show that performance is considerably increased as the generalised dataset is loaded 2.8 times faster. Similarly Ellul and Altenbuchner (2014) examine the impact of generalisation on rendering of 3D city models in mobile devices. Reducing the data volume by generalisation operators (aggregation and simplification) as well as eliminating duplicate faces by topological data structuring significantly increases the 3D model rendering performance with results on two mobile devices showing 7.06 times and 9.54 times faster performance.
Biljecki et al. (2016) present an experiment where changes in modelling choices have significant different effects when employing those models in spatial analyses. They create a range of different building models in LoD1 and LoD2 based in different geometric references and they investigate three different spatial analyses with each model: area of the building envelope, volume of the building and shadow casted by a building. They conclude that there is no optimal geometric structure suitable for multiple analysis processes and that errors and their distribution are caused by alterations in the configuration of buildings which makes them unique for each geometric reference (ibid).

The above review highlights the fact that, to date, relatively little work has been carried out to determine not only the performance impact of generalisation but the overall impact on algorithm results. In other words, while performance may be greatly improved in terms of time to execute an algorithm, is this at the cost of the 'fitness for purpose' of the results? The remainder of this paper describes first tests to assess this issue, for 3D analytical algorithms of sun shadow volume calculation and line of sight.

3. DATA

The dataset used for the experiment is the 'Ordnance Survey MasterMap Topography Layer, including Building Height Attribute'[1] information supplied by the centre for digital expertise of the University of Edinburgh EDINA through their web mapping portal Digimap[2] (© Crown Copyright and Database Right 2018. Ordnance Survey - Digimap Licence). It provides height attributes to the buildings within the OS MasterMap Topography Layer obtained from the OS Digital Terrain Model[3], using Ordnance Datum Newlyn as the national height datum within the OSGM15 National Geoid Model and British National Grid spatial reference system. The heights used in this experiment is the difference height value between the absolute ground level and the highest part of the roof of a building.

An absolute accuracy of 0.9 m and a relative accuracy of +/- 1.1 m is applicable to the horizontal accuracy of the dataset as it is based on the 1:1250 OS MasterMap4. In regards of the vertical accuracy, this is again dependent to the source of the data used to obtain the vertical values, which in this case are the DSM and DTM. Because this dataset is still in a BETA state and subject to update, confidence levels are provided instead of accuracy values.

The study area is a 5 km by 5 km dataset of buildings from Greenwich and Canary Wharf in South East London which has been selected due to its mix of low level suburban housing and taller buildings (Figure 1).
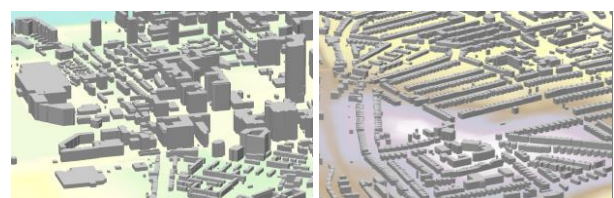

Figure 1. 3D Buildings of different heights[4]

---

[1] https://www.ordnancesurvey.co.uk/docs/product-guides/osmm-topo-layer-bha-product-guide.pdf
[2] https://digimap.edina.ac.uk/os

[3] https://ordnancesurvey.co.uk/business-and-government/products/terrain-50.html
[4] https://www.ordnancesurvey.co.uk/business-and-government/help-and-support/products/topography-layer.html

4. METHODOLOGY

For this experiment the following off-the-shelf software is used: ESRI's ArcGIS 10.5 (ESRI, 2008), FME 2017.1 (Safe Software, 2018) and pgAdmin 4 v2 (The PostgreSQL Global Development Group, 2018) alternating them in order to satisfy the flaws encountered in each software on a PC Intel(R) Core(TM) i7-7600U CPU 2.80 GHz, 16.0 GB RAM on a Window 10 64-bit.

4.1 Generalisation and extrusion – Full Dataset

The generalisation process applied to the original data (5 km by 5 km) is composed by an aggregation of buildings, followed by a simplification process. ESRI's ArcGIS software was used for both operators. Extrusion of the results in ArcScene[5] is held prior to the spatial analyses.

4.1.1 Aggregation

The first step is the aggregation[6] of the raw data with a tolerance of 1 m, a minimum size of 25 m^2 for every building and an internal minimum hole size of 25 m^2. When inspecting the resulting dataset, a flaw with this tool was detected: while the buildings are correctly aggregated, holes smaller than 25 m^2 are retained. This issue is resolved by repeating the aggregation process, as shown on Figure 2. For each aggregated polygon, the average height is calculated by summing the height values for each building and dividing by the number of buildings.
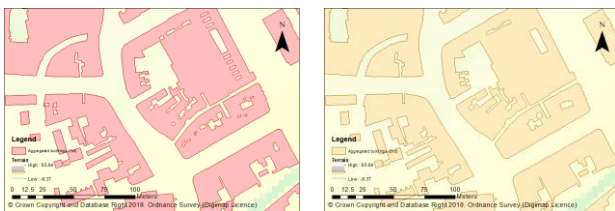

Figure 2. First aggregated buildings (left) and hole fill-up (right)

Both original and aggregated buildings are translated using SafeSoftware's 'FME Data Inspector' to a PostgreSQL database using the PostGIS extension in order to calculate the average height of the resulting building blocks.

4.1.2 Simplification

The building polygons obtained from the second round of aggregation process are simplified by ESRI's tool 'simplification of buildings'[7] using a minimum building area of 25 m^2 and a tolerance of 1 m and 5 m (Figure 3).

Figure 3. Simplified data (tolerances of 1 m left and 5 m right)

4.1.3 Extrusion and visualisation

Simplification completes the 2D generalisation process. 3D visualisation of the original data is obtained by extruding[8] the building polygons to their height attribute and generalised datasets to their average height attribute calculated in steps above (Section 4.1.1) using ESRI's ArcScene software.

4.2 Reducing the Dataset Size

Once three-dimensional buildings are obtained by extrusion, the impact of the generalisation can be investigated in the selected use cases. However, in the first trial of the spatial analysis, the chosen ESRI's tool could not deliver any result due to the large amount of data (57999 polygons) and the characteristics of the computer used. As a solution to the issue, a reduced 1 km by 1 km area is selected maintaining the same essence as the original dataset: a mix of buildings with different heights.

Figure 4 shows a 2D maps of the buildings in the 1 km by 1 km reduced dataset in the original situation (a), data aggregated using a 1 m tolerance (b), simplified using a 1 m tolerance (c) and simplified using a 5 m tolerance (d).
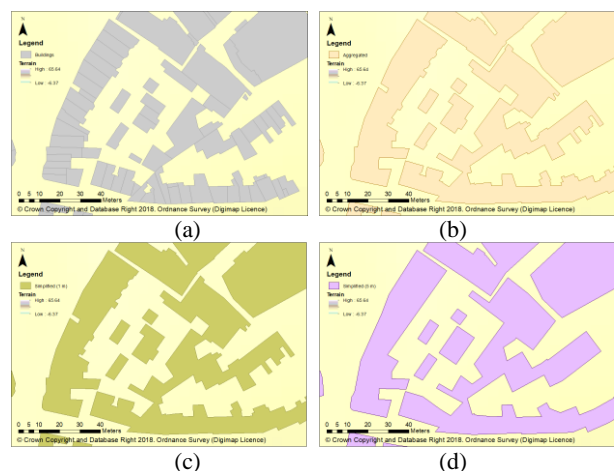

Figure 4. Original and generalised datasets in a larger scale

The three-dimensional visualisation of the original data is carried out as above (Section 4.1.3).

4.3 Accuracy measurements

In order to determine the accuracy of the information derived from the generalised dataset when compared to the original model, four measurements are proposed as indicators of change: centroid shift, area, perimeter and volume change.

4.3.1 Centroid shift

The centroid of every polygon is calculated in ArcGIS for the original - $C_o(X_o, Y_o)$ and generalised - $C_g(X_g, Y_g)$ dataset and translated into a PostgreSQL database using FME. In PostGIS, the function 'ST_Intersects'[9] is used to evaluate which original polygons have been aggregated into each generalised polygon.

---

[5] http://desktop.arcgis.com/en/arcmap/10.5/extensions/3d-analyst/the-arcscene-user-interface.htm

[6] http://desktop.arcgis.com/en/arcmap/10.5/tools/cartography-toolbox/aggregate-polygons.htm

[7] http://desktop.arcgis.com/en/arcmap/10.5/tools/coverage-toolbox/simplify-building.htm

[8] http://desktop.arcgis.com/en/arcmap/latest/extensions/3d-analyst/about-using-extrusion-as-3d-symbology.htm

[9] http://postgis.net/docs/ST_Intersects.html

(a)   (b)

Figure 5. Centroid change control.

The overall coordinates of the geometric centre of the centroids on those original polygons (Figure 5a) are computed with the function 'ST_Centroid'[10] averaging the X and Y coordinates of the vertices involved:

$$C_c(X_c, Y_c) = \left(\frac{X_{O1} + X_{O2} + \cdots + X_{On}}{n}, \frac{Y_{O1} + Y_{O2} + \cdots + Y_{On}}{n} \right) \qquad (1)$$

where C_C = calculated geometric centre of original centroids
X_C, Y_C = coordinates of calculated geometric centre
X_O, Y_O = coordinates of the original centroids
n = number of original centroids involved

The outcomes are translated using FME into a shapefile in order to calculate the difference in X and Y coordinates between the calculated centroid of the original polygons and that of the resulting aggregated polygon (Figure 5b), and therefore, the distance between them as:

$$D_G^C = \sqrt{(\Delta X_G^C)^2 + (\Delta Y_G^C)^2} \qquad (2)$$
$$\Delta X_G^C = X_G - X_C$$
$$\Delta Y_G^C = Y_G - Y_C$$

where D_G^C = Distance between centroids
ΔX = Centroid difference in X axis
ΔY = Centroid difference in Y axis
X_C, Y_C = coordinates of calculated geometric centre
X_G, Y_G = coordinates of the centroid on the generalised dataset

Overall statistics were then calculated for the values obtained, and specific cases showing significantly large shifts explored in more detail.

4.3.2   Area, perimeter and node count change
A similar procedure is carried out to evaluate changes in area, perimeter and nodes of the real buildings compared to the generalised buildings. The area, perimeter, and node counts are obtained as:

$$A_c = A_{O1} + A_{O2} + \cdots + A_{On} \qquad (3)$$
$$P_c = P_{O1} + P_{O2} + \cdots + P_{On}$$
$$N_c = N_{O1} + N_{O2} + \cdots + N_{On}$$

where A_C, P_C, N_C = Calculated area, perimeter and number of nodes from the original polygons involved.
A_O, P_O, N_O = Area, perimeter and number of nodes of each polygon on the original dataset
n = number of original polygons involved

The differences are obtained as:
$$\Delta A_G^C = A_G - A_C \qquad (4)$$

---

[10] http://postgis.net/docs/ST_Centroid.html
[11] http://desktop.arcgis.com/en/arcmap/10.3/tools/3d-analyst-toolbox/sun-shadow-volume.htm
[12] http://desktop.arcgis.com/en/arcmap/10.3/tools/3d-analyst-toolbox/add-z-information.htm

---

$$\Delta P_G^C = P_G - P_C$$
$$\Delta N_G^C = N_G - N_C$$

where A_G, P_G, N_G = Area, perimeter and number of nodes of the polygon on the generalised dataset
ΔA = Area difference
ΔP = Perimeter difference
ΔN = Number of nodes difference

4.4   Sun Shadow

The first spatial analysis carried out with the finalised 1 km by 1 km extension 3D models is the shadow casting of each individual building by using the ArcScene tool 'Sun Shadow Volume'[11] on both raw (Figure 6a) and generalised (Figure 6b) datasets. The date and time fixed is 2018-05-15 at 15:30.

(a)   (b)

Figure 6. Sun shadow casting from obtained 3D models.

Run time of the analysis is recorded in order to evaluate the impact of the generalisation. In addition, the volume of each building is calculated with the tool 'Add Z information'[12] which obtains the total sum of volume of the original and generalised 3D models. The differences between the two scenarios are calculated as:

$$\Delta V_G^O = V_G - V_O \qquad (5)$$
$$\Delta ts_G^O = ts_G - ts_O$$

where V_O = Total volume of the resulting shadows from the original 3D model
V_G = Total volume of the resulting shadows from the generalised 3D model
ts_O = Analysis duration for the original 3D model
ts_G = Analysis duration for the generalised 3D model
ΔV = Total volume difference
Δts = Time/duration difference

4.5   Line of Sight

The original and generalised 3D models are used for the second spatial analysis where a hypothetical mobile antenna mast is installed on the roof of a prominent building. The position of the antenna in both models is the same: XYZ (537957.000, 177630.000, 67.100). Sight lines from the antenna are first determined by using the ArcScene tool 'Construct Sight Lines'[13] and the visibility from those is defined by 'Line Of Sight'[14] (Figure 7). In addition, a point feature class is created with the information from where the line of sight has been first obstructed against the 3D models.

---

[13] http://desktop.arcgis.com/en/arcmap/10.3/tools/3d-analyst-toolbox/construct-sight-lines.htm
[14] http://desktop.arcgis.com/en/arcmap/10.3/tools/3d-analyst-toolbox/line-of-sight.htm
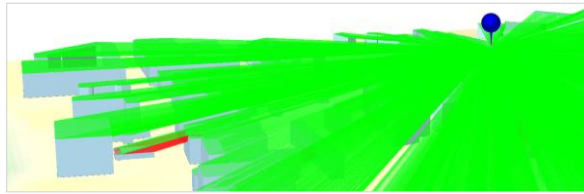

Figure 7. Sight lines from the antenna: visible (green) and non-visible (red).

The geometry enclosing all those points is generated, in other words, a convex hull and its area determines the visible surface from the observation point. The interval the software needs to explore the spatial analysis is recorded in order to see the performance:

$$\Delta Ac_G^O = Ac_G - Ac_O \qquad (6)$$
$$\Delta tl_G^O = tl_G - tl_O$$

where Ac_O = Convex hull area on the original 3D model
           Ac_G = Convex hull area on the generalised 3D model
           tl_O = Analysis duration for the original 3D model
           tl_G = Analysis duration for the generalised 3D model
           ΔAc = Total area difference
           Δtl = Time/duration difference

5. RESULTS

5.1 Generalisation results – Full Dataset

The geographic information of the original 5 km by 5 km dataset has been reduced significantly, i.e. 57999 buildings were generalised into 9879 buildings. Table 2 summarises the values for each new dataset obtained throughout the generalisation steps.

Dataset (5 km by 5 km)	No. of Polygons	No. of Nodes	Area (m²)	Volume (m³)
Original (raw)	57999	508009	4513848.959	70230714.429
Aggregated (1st)	9887	304382	4484212.351	N/A
Aggregated (2nd)	9883	302812	4486751.712	59122900.093
Simplified (1 m)	9883	188365	4522861.667	59200040.927
Simplified (5 m)	9879	83042	4582619.579	59790073.082

Table 2. Obtained values throughout the generalisation process to the total original 5 km by 5 km dataset.

The outcome of comparing each generated dataset with the original data is presented in percentages in Table 3 in order to further examine the impact of the generalisation process. Negative values represent a decreases in information while the positive values mean an increase, e.g. comparing the total area of all buildings within the original dataset with the total area of the most generalised dataset results in an increase of 1.52 %, due to their replacement with larger polygons after simplification.

Polygons and node counts are reduced by approximate 83% while the total volume (obtained from the sum of volume of every single building in 3D) is reduced by around 15% due to the change on the height of the buildings on the aggregation step.

Differences (5 km by 5 km)	No. of Polygons	No. of Nodes	Area (m²)	Volume (m³)
Aggregated (2nd)	-82.96 %	-40.39 %	-0.60 %	-15.82 %
Simplified (1 m)	-82.96 %	-62.92 %	0.20 %	-15.71 %
Simplified (5 m)	-82.97 %	-83.65 %	1.52 %	-14.87 %

Table 3. Reduction/addition of data in percentages after generalisation to the total original 5 km by 5 km dataset.

5.2 Generalisation results – Reduced Dataset

As the 5 km by 5 km dataset is minimized into a 1 km by 1 km in order to test the impact of the generalisation when realising some spatial analyses, new results from this smaller version of raw data were obtained in every generalisation operator (see Table 4).

Dataset (1 km by 1 km)	No. of Polygons	No. of Nodes	Area (m2)	Volume (m3)
Original (raw)	2255	22969	236206.223	3669342.439
Aggregated (1st)	410	13493	235342.000	N/A
Aggregated (2nd)	410	13440	235406.520	3361866.650
Simplified (1 m)	410	8155	235400.327	3361548.335
Simplified (5 m)	410	3878	236953.666	3378923.765

Table 4. Obtained values throughout the generalisation process to the total minimized 1 km by 1 km dataset.

Taking an identical approach to that used for the whole dataset, the results obtained in every step are compared with the new reduced dataset and shown as percentages in Table 5. Similarly, to the whole 5 km by 5 km dataset, polygons and nodes are reduced by circa 82% while there is less impact on area and volume.

Differences (1 km by 1 km)	No. of Polygons	No. of Nodes	Area (m²)	Volume (m³)
Aggregated (2nd)	-81.82 %	-41.26 %	-0.37 %	-8.38 %
Simplified (1 m)	-81.82 %	-64.50 %	-0.34 %	-8.39 %
Simplified (5 m)	-81.82 %	-83.12 %	0.32 %	-7.91 %

Table 5. Reduction/addition of data in percentages after generalisation to the total original 1 km by 1 km dataset.

5.3 Accuracy measurement results

One of the most interesting factors which determines how much the geographic information suffered a change is the distance between the geometric centre of the original centroids and the centroids on the generalised dataset (for the whole 5 km by 5 km study area). The shift of these centroids has been plotted as the frequency of the values obtained for both the differences of the coordinates in X, ΔX (Figure 8a) and in Y, ΔY (Figure 8b) on a logarithmic log-10 scale.
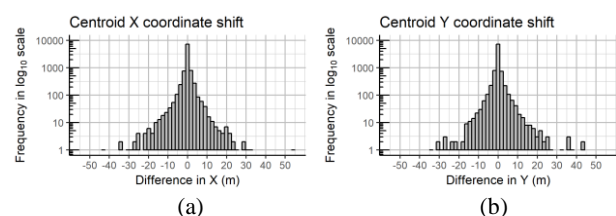

Figure 8. Histograms representing the frequency of the observations appearance on *y* axis for ΔX (a) and ΔY (b) shifts on *x* axis, on a log-10 scale. For an easier visual interpretation, the frequencies have been split in 35 bins of 1.8 bin width.

Examining an extreme case, Figure 9 displays the calculated geometric centre from the original polygons' (no. 3012 - left) and the generalised polygon and its centroid (no. 5912 - right). Visualisation of both together (right) evidences the existence of a large centroid shift, the maximum between all in this case.
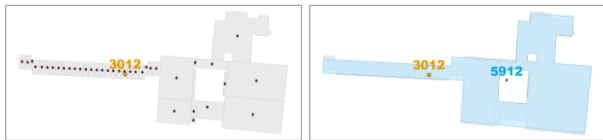
Figure 9. Centroid maximum shift case.

In addition to the centroid shift, statistical results for the remaining measures (area, perimeter and number of nodes change) are presented in Table 6.

Differences	Minimum	Maximum	Mean	Standard Deviation
ΔX (m)	-42.825	53.571	-0.033	2.758
ΔY (m)	-34.462	43.098	0.024	2.763
D_G^C (m)	0	53.688	1.704	3.510
ΔA (m^2)	-7605.056	572.058	5.402	82.433
ΔA (%)	-78.06	23.07	0.76	4.22
ΔP (m)	-8228.524	6.259	-98.288	184.261
ΔP (%)	-83.23	20.592	-36.206	21.094
ΔN (points)	-2726	0	-41.311	71.334
ΔN (%)	-98.76	0	-69.27	23.84

Table 6. Statistical analysis- Measures of Change.

5.4   Sun Shadow results

At a glance, the generated sun shadow volumes look visually correct. However, navigating throughout the three-dimensional space a flaw in the tool is detected: shadow casting continues underneath terrain (Figure 10). Consequently, the obtained volume values are larger than expected as they include the volume below ground. As this issue is consistent in all tests, it is still relevant to measure overall volume change due to generalisation as all the volume measurements will include the extra volume.
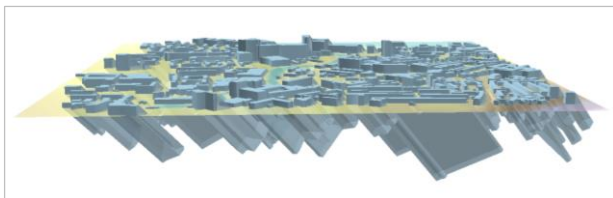

Figure 10. Shadow casting belowground.

Results show that the shadow volume is reduced by 11.58% (ΔV = - 4176325.31 m^3) from the original (V_O = 36060383.30 m^3) to the generalised 3D model (V_G = 31884057.99 m^3). Similarly, an important 6.32 times (Δts = - 44.99 s) improvement in calculation results from the reduction in detail (ts_O = 53.44 s, ts_G = 8.45 s).

5.5   Line of Sight results

Using the generalised dataset, the total area where the visibility of sight lines is not interrupted is increased by 19.81% (ΔAc = 32641.47 m^2) from the original 3D model (Ac_O = 164792.26 m^2) to the generalised one (Ac_G = 197433.73 m^2). The process runs 7.82 times faster (Δtl = - 10651.29 s) on the generalised model (tl_G = 1560.96 s) than the original model (tl_O = 12212.25 s).

6.   DISCUSSION

This paper addresses the question: *What is the impact of generalisation on the results of 3D line-of-sight and shadow-*

casting algorithms? From the results it can be seen that applying 2D generalisation operators to a dataset prior to extruding it into a 3D model involves a loss of detail in terms of the model (14.87% less volume) while performance in the spatial analyses algorithms is improved significantly: 7.82 times faster for the line of sight and 6.32 times faster when casting shadows. The overall increase in footprint (area) results in a decrease in shadow volume of over 11%, which could be fairly significant when estimating, for example, the potential for solar panel energy across a city. Similarly, generalised data resulted in a nearly 20% increase in visible areas for line of sight for a hypothetical mobile mast. Given the importance of the accuracy of such calculations – in particular with emerging 5G technology – this suggests that mobile phone companies should use detailed data where possible, even though this comes at greater computational cost and potentially at greater purchase cost.

This study reflects previous findings by Ellul and Joubran (2012) and highlights the potential of displaying larger 3D datasets and using them more efficiently for comprehensive 3D spatial analyses. This can be particularly useful in situations where detailed representation is not needed or where more detailed data is not available. However, whether using generalised or original data for each spatial analysis depends on the user case scenario and its detail requirements, which confirms findings by Biljecki et al. (2016) that one size does not fit all when it comes to generalisation.

Issues encountered when using off-the-shelf tools for the spatial analyses (specifically problems when running line of sight using a large dataset and the casting of shadows belowground) demonstrate that further work is still required before such tools can be considered fully mature. This may currently be one factor that is limiting the uptake of 3D GIS specially when working with larger datasets. The study also highlighted the lack of commercially available 3D generalisation tools, and while 2D generalisation with extrusion was used as a workaround, the results obtained were consequently limited to the LOD1 data – i.e. flat roofs.   As this approach also involved an average height for a block, local over-and-under estimations in height could be significant.

7.   CONCLUSION AND FURTHER WORK

The work carried out above highlights the importance of 3D generalisation, in particular as newer 3D analysis algorithms come on stream; understanding the impact of detail versus generalised – and which components of a model should be kept and which could be generalised in each application case is important – both for performance and for analysis.

As more extensive and more detailed datasets become available – e.g. OS has introduced height information for all buildings in GB - the presented exploration of 3D generalisation is particularly important when up and coming requirements for 3D are considered –specifically smart cities, and the need to locate sensors in 3D space and perform analysis based on their location. For instance, the impact of line of sight is of special interest for 5G telephony.

Further work is suggested in order to understand different user case needs in depth, and in their respective application domains. Technically, 3D generalisation is not a solved problem, and there is much room for improvement even within the above approach - for example, it may be possible to take the height of individual buildings into account when creating the block – and if the height

is over a certain percentage above the mean – such as a church tower – then keep the original building.

Limitations with the software used suggest need for further work in order to generate robust tools that will in turn encourage uptake of 3D analysis.

ACKNOWLEDGEMENTS

REFERENCES

Baig, S.U., Rahman, A.A., 2013. Generalization and Visualization of 3D Building Models in CityGML. In: Pouliot, J., Daniel, S., Hubert, F., Zamyadi, A. (Eds.), *Progress and New Trends in 3D Geoinformation Sciences*. Springer, Berlin, Heidelberg, pp. 63–77, doi.org/10.1007/978-3-642-29793-9_4.

Biljecki, F., Ledoux, H., Stoter, J., Vosselman, G., 2016. The variants of an LOD of a 3D building model and their influence on spatial analyses. In: *ISPRS J. Photogramm. Rem. Sens.*, Vol. 116, pp. 42–54, doi.org/10.1016/j.isprsjprs.2016.03.003.

Biljecki, F., Stoter, J., Ledoux, H., Zlatanova, S., Çöltekin, A., 2015. Applications of 3D City Models: State of the Art Review. In: *ISPRS Int. J. Geo-Inform.*, Vol. 4(4), pp. 2842–2889, doi.org/10.3390/ijgi4042842.

Christen, M., Nebiker, S., 2015. Visualisation of complex 3D city models on mobile webbrowsers using cloud-based image provisioning. In: *ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences,* Vol. II-3/W5, pp. 517–522, doi.org/10.5194/isprsannals-II-3-W5-517-2015.

Ellul, C., Altenbuchner, J., 2014. Investigating approaches to improving rendering performance of 3D city models on mobile devices. In: *Geo-spatial Information Science*, Vol. 17(2), pp. 73–84, doi.org/10.1080/10095020.2013.866620.

Ellul, C., Joubran, J.A., 2012. Preliminary investigations into the potential of improving rendering performance of 3D datasets using 2D generalisation. In: *Usage, Usability, Util. 3D City Model,* 03005, doi.org/10.1051/3u3d/201203005.

ESRI, 2018. About ArcGIS Mapping And Analytics Platform http://www.esri.com/en-us/arcgis/about-arcgis/overview (13 July 2018).

Fan, H., Meng, L., 2012. A three-step approach of simplifying 3D buildings modeled by CityGML. In: *International Journal of Geographical Information Science*, Vol. 26, pp. 1091–1107, doi.org/10.1080/13658816.2011.625947.

Fan, H., Meng, L., Jahnke, M., 2009. Generalization of 3D Buildings Modelled by CityGML, In: Sester, M., Bernard, L., Paelke, V. (Eds.), *Advances in GIScience, Lecture Notes in Geoinformation and Cartography*. Springer Berlin Heidelberg, Berlin, Heidelberg, pp. 387–405, doi.org/10.1007/978-3-642-00318-9_20.

Forberg, A., 2007. Generalization of 3D building data based on a scale-space approach. In: *ISPRS Journal of Photogrammetry and Remote Sensing,* Vol. 62(2), pp. 104–111, doi.org/10.1016/j.isprsjprs.2007.01.002.

Glander, T., Döllner, J., 2009. Abstract representations for interactive visualization of virtual 3D city models. In: *Computers, Environment and Urban Systems*, Vol. 33(5), pp. 375–387, doi.org/10.1016/j.compenvurbsys.2009.07.003

Gröger, G., Kolbe, T., Nagel, C., Häfele, K.-H., 2012. OGC City Geography Markup Language (CityGML) En-coding Standard, Open Geospatial Consortium. https://portal.opengeospatial.org/files/?artifact_id=47842.

Herbert, G., Chen, X., 2015. A comparison of usefulness of 2D and 3D representations of urban planning. In: *Cartography and Geographic Information Science*, Vol. 42, pp. 22–32, doi.org/10.1080/15230406.2014.987694.

Hofierka, J., Zlocha, M., 2012. A New 3-D Solar Radiation Model for 3-D City Models. In: *Transactions in GIS*, Vol. 16(5), pp. 681–690, doi.org/10.1111/j.1467-9671.2012.01337.x.

Kada, M., 2002. Automatic Generalization of 3D Building Models. In: *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, Vol. XXXIV Part 4, pp. 243–248.

Kemec, S., Duzgun, S., Zlatanova, S., Dilmen, D.I., Yalciner, A.C., 2010. Selecting 3D urban visualisation models for disaster management: Fethiye tsunami inundation case. In: *Proceedings of the 3rd International Conference on Cartography and GIS.*

Lafarge, F., Mallet, C., 2012. Creating Large-Scale City Models from 3D-Point Clouds: A Robust Approach with Hybrid Representation. In: *International Journal of Computer Vision*, Vol. 99(1), pp. 69–85, doi.org/10.1007/s11263-012-0517-8.

Ledoux, H., Meijers, M., 2011. Topologically consistent 3D city models obtained by extrusion. In: *International Journal of Geographical Information Science*, Vol. 25(4), pp. 557–574, doi.org/10.1080/13658811003623277.

Robinson, A.H., Morrison, J.L., Muehrcke, P.C., Kimerling, A.J., Guptill, S.C., 1995. Elements of Cartography, 6th ed., New York, pp. 449-471.

Safe Software, 2018. FME Desktop - Data Integration And Productivity https://www.safe.com/fme/fme-desktop/ (13 July 2018).

Sester, M., Brenner, C., 2005. Continuous Generalization for Visualization on Small Mobile Devices, In: *Developments in Spatial Data Handling*. Springer, Berlin, Heidelberg, pp. 355–368, doi.org/10.1007/3-540-26772-7_27.

Shojaei, D., Olfat, H., Briffa, M., Rajabifard, A., 2017. 3D Digital Cadastre Journey in Victoria, Australia. In: *ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, Vol. IV-4/W5, pp. 117–123, doi.org/10.5194/isprs-annals-IV-4-W5-117-2017.

Stoter, J., De Kluijver, H., Kurakula, V., 2008. 3D noise mapping in urban areas. In: *International Journal of Geographical Information Science*, Vol. 22(8), pp. 907–924, doi.org/10.1080/13658810701739039.

Stoter, J.E., Ploeger, H.D., 2003. Property in 3D—registration of multiple use of space: current practice in Holland and the need for a 3D cadastre. In: *Computers, Environment and Urban Systems*, Vol. 27(6), pp. 553–570, doi.org/10.1016/S0198-9715(03)00014-0.

The PostgreSQL Global Development Group, 2018. https://www.postgresql.org/ftp/pgadmin/pgadmin4/v2.0/windows/ (4 April 2018).