ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences, Volume IV-4/W6, 2018
13th 3D GeoInfo Conference, 1–2 October 2018, Delft, The Netherlands

USER REQUIREMENTS GATHERING FOR A NATIONAL 3D MAPPING PRODUCT IN THE UNITED KINGDOM

K. Wong[a] and C. Ellul[b]

[a] Department of Computer Science, University College London, UK - kelvin.wong.11@ucl.ac.uk
[b] Department of Civil, Environmental and Geomatic Engineering, University College London, UK - c.ellul@ucl.ac.uk

KEY WORDS: 3D GIS; 3D geoinformation; NMCAs; user requirements; United Kingdom.

ABSTRACT:

One of the main barriers to the uptake of 3D GI is the lack of understanding of what the user requirements are. From the data acquisition and creation perspective – in particular, that of a National Mapping and Cadastral Agency who may need to prepare datasets with national coverage – this is an issue as each new 3D feature type and element within a feature added (such as doors, windows, chimneys, street lights) requires additional processing and cost to create. This paper reports the results of a user requirements gathering exercise for a national 3D mapping product in the United Kingdom. The study focuses on the user perception of 'usefulness' of different 3D geometry and semantic features. A web-based questionnaire with Likert-type items was selected as the primary data collection method and was conducted in May 2017. A total of 121 completed responses were from the UK. Descriptive analysis showed that 'Air quality engineering', 'Infrastructure & transport' and 'Environmental services' presented the most positive outlook on the usefulness of 3D. Correlation analysis showed that potential 3D product groups that could be formed in a multi-product approach. Cluster analysis showed that appetite for 3D information not only varies between sectors, but also within sectors between different practitioners. The results from exploratory factor analysis showed that users were more interested in additional information on non-building features rather than additional detail to building geometry. Further continued work is required to incorporate both non-GIS users and 3D users outside the GIS domain.

1. INTRODUCTION

One of the main barriers to the uptake of 3D is the lack of understanding of what the user requirements are (Stoter et al., 2013), a question of particular relevance to a National Mapping and Cadastral Agency (NMCA), where national coverage datasets might be considered. While requirements for 2D geographic information have evolved over time through NMCAs working iteratively with their end users over many years, 3D geographic information requirements, are relatively nascent and this iterative process of user requirements development has not yet taken place.

The relative lack of clear requirements offers a high degree of freedom to define what 3D geographic information is, from the data acquisition perspective, each new 3D feature type and element within a feature added (such as doors, windows, chimneys, street lights) requires additional processing and cost to create. As such, it is relevant to **understand the importance of different 3D features for different applications**, allowing data producers to prioritise features to be captured in 3D. A second question to be addressed **is whether a traditional 'one-size-fits-all' approach is appropriate, or if multiple tailored products are more useful for the user**. From the perspective of the data producer, a single product would result in lower overall costs in production and maintenance. Multiple products would lead to increased costs, although this is ameliorated if they are produced from the same source data. Thus, it is also useful to identify whether, if a user finds one type of 3D feature useful, then do they tend to find another type of 3D feature useful as well and hence identify potential 3D product groups that could be defined to satisfy the user needs.

The study explores both issues by firstly issuing a questionnaire to determine 3D features of interest to the different domains, and then investigating the potential of classifying applications demonstrating similar 3D requirements together to produce the potential product groups, working towards the bigger end goal of answering "what should a national 3D dataset look like in the UK?". The work described in this paper forms a part of a larger project of developing a national 3D mapping product in collaboration with Ordnance Survey, the national mapping agency of Great Britain.

2. LITERATURE REVIEW

2.1 Potential applications of 3D GI

Applications of 3D GI are widely varied and include support for mineral discovery, noise mapping, public rescue operations, ecological studies, and utility management, as well as assessing propagation and impact of air pollutants to protecting city skylines to creating digital historic dioramas. To date, the most cited applications focused on visualisation or analyses that could be conducted using datasets containing simple representations of building geometry. These include, but are not limited to calculating solar potential, estimating flood potential, estimating noise propagation, calculating viewsheds and shadow analysis. Thus, current use of 3D GI is dominated by visualisation-based applications with a focus on building geometry. While 3D visualisations are used to support decision-making processes and to allow for better communication, but there are few instances of literature citing the benefits of 3D analyses. One possible explanation is the fact that 3D is still a maturing technology (despite the many years of development) and there is still a lack of software, hardware, data, and expertise despite the favourable trends in the last 5 years.

2.2 Production of 3D geographic information

There are many different methods to produce 3D geographic information, at different coverages, scales, accuracies, and levels of detail. Basic 'block' models with flat roofs can be easily created by extruding 2D building footprints with height information derived from LiDAR or photogrammetric surveys. For more detailed buildings, these can be generated using Computer Aided Design (CAD) data or Building Information Models (BIM). One issue here, however, is that making the models compatible for GI systems can be error-prone as the schemas between the different 3D formats do not always map

This contribution has been peer-reviewed. The double-blind peer-review was conducted on the basis of the full paper.
https://doi.org/10.5194/isprs-annals-IV-4-W6-89-2018 | © Authors 2018. CC BY 4.0 License.

one-to-one. The production also tends to be very manual, time-consuming, tedious, subjective and requiring skill (Tang et al., 2010). Another method is the use of mobile mapping systems fitted with photogrammetric and LiDAR systems. Mobile vehicles are used to map features from the ground-level, providing better results for vertical surfaces such as building facades, in comparison to those obtained from oblique airborne data sources. Lastly, one consideration during the production of 3D geographic information is the inclusion of textures (also known as texturing or texture mapping). Texturing refers to the application of an image to the surface of a polygon and can be derived photogrammetrically or produced procedurally. The inclusion of textures has the ability to enhance the photorealism of 3D visualisations but can cause performance issues where high-resolution 'real' imagery is used (as each surface contains a unique image).

2.3 3D Data requirements gathering

Several studies have conducted different forms of requirements analysis of 3D data (Biljecki et al., 2015; Sargent et al., 2015; Stoter et al., 2013; Stoter et al., 2016; Walter, 2014; Wong and Ellul, 2017) - more details on these studies can be found in Wong and Ellul (2017). Two key points emerge from these studies. Firstly, user requirements for 3D GI are application specific. For example, while roof geometry may be important for estimating solar potential of buildings, it is less important for urban pedestrian navigation applications (as the roof of tall buildings often cannot be seen). Secondly, user requirements for 3D GI are country specific. For example, in Finland where over 75% of land area is forested (Finnish Forest Association, 2016), the use of 3D within forestry management allows for dramatically reduced survey costs, reduced logistical costs and increased forest productivity (Tuokko, 2017). In contrast, only 13% of the United Kingdom is wooded (Forestry Commission, 2017) and there is less of a driver for 3D in forestry applications. In another example, many countries (such as Denmark and Australia) have a national cadastre for land administration. In these cases, an extension to 3D can provide more uniform assessment of payable land tax and improve information for notaries thus speeding up transaction time and lowering associated costs (Witmer, 2017). The United Kingdom, however, does not have a national cadastre. Instead, land laws are based on 'estates' and the concept of rights through time rather than simple ownership. As such, the extension to a 3D cadastral system in the UK would be unfeasible. In summary, it is therefore important to explore user requirements for 3D GI at the application-specific and the country-specific level, to allow NMCAs and other data producers to prioritise work towards the needs of the national audience. Within the context of the UK, a study by Wong and Ellul (2017) explored user requirements gathering for 3D geographic information using qualitative methods. Web-based questionnaires and semi-structured face-to-face interviews were used to assess the state of 3D GI adoption and use within the UK.

3. METHODOLOGY

3.1 Data collection

A web-based questionnaire with Likert-type items was selected as the primary data collection method in this study. In previous web-based questionnaires on 3D user requirements (Wong and Ellul, 2017), there were poor response rates when participants were presented with open-ended questions which required an original and personal response. To increase the response rate, the 'Usefulness of 3D[1]' questionnaire was specifically designed to be very short and easy to answer. The web-based questionnaire utilised five-point Likert-type items[2] to gauge the participants' perception on the usefulness of different 3D information. In particular, the main question of interest was:

> *Please rate the usefulness of the following 3D information according to your day-to-day work*

The participants were asked to rate each suggested 3D information on a unipolar scale: 5 = 'Extremely useful'; 4 = 'Very useful'; 3 = 'Moderately useful'; 2 = 'Slightly useful'; 1 = 'Not at all useful'; or 'Not applicable'. Representative images were also included to aid the user in understand each item. By presenting a rating scale over a dichotomous question, it captured the participant's prioritisation of 3D GI requirements. The list of 3D geographic information included was derived from a combination of reviewing applications within existing literature (Section 2.1) as well as discussions with Ordnance Survey on what potential 3D information could be produced. It is acknowledged that this is not an exhaustive list, but provides an indication of the most common 3D information. In addition to rating the usefulness of 3D information, participants were also asked to briefly describe the specific task where they use 3D GI if they had selected 'Extremely useful'. Participants were asked for their organisation's name and the sector they worked in (with the list of sectors being derived from literature – see Section 2.1), for context, with multiple options being allowed here to reflect the interdisciplinary nature of GIS. An excerpt of the questionnaire is depicted below in Figure 1. As part of the questionnaire design, a sixth 'Not Applicable' category was also included for each question. This was because the participants were from a wide range of sectors and not all datasets were related to their field of work. By having a 'Not Applicable' category allowed participants to provide a reply to every question without forcing a response. Within the subsequent analysis, 'Not Applicable' responses were coded as missing data to avoid introducing additional bias e.g. coding 'Not Applicable' as the value '0'.

3.2 Participants

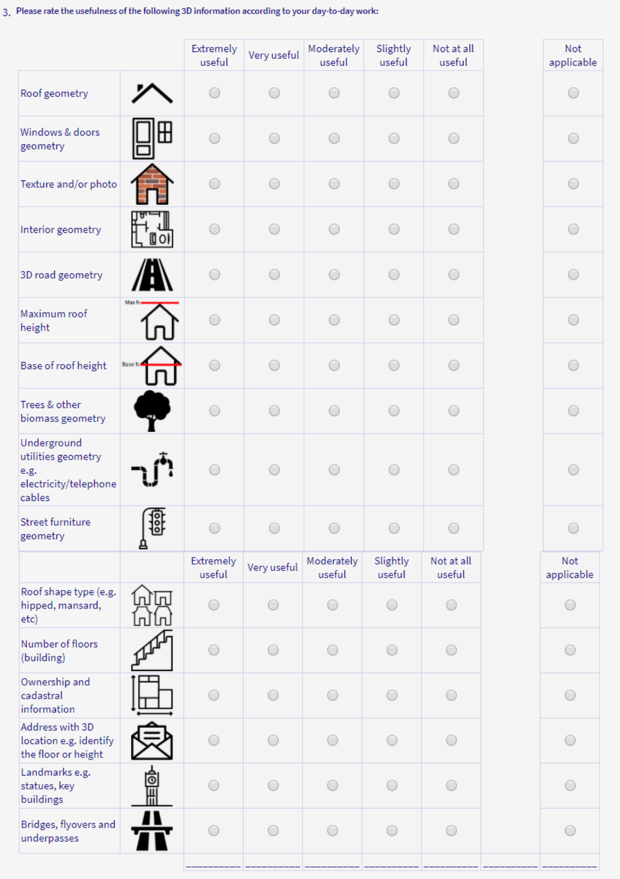
The main target group of the questionnaire and interviews were professionals who work directly with geographic information. GIS practitioners were targeted for their expertise as they were most likely to have the best understanding of how geographic information was used in their organisation. There is also a working assumption that these existing users of 2D GI were most likely to become the early adopters of 3D GI. It is important to note, however, the possible bias of the community

---

[1] 'Usefulness' can be a subjective term - in this paper, it is defined as the quality of a product, dataset or solution to achieve a user's goal or goals. In essence, it is a user's perception of the fitness-for-purpose of a product, acting as an indicator of 'value' of a product as ascribed by the user. By understanding what is perceived to be 'useful', data producers can begin to design and tailor 3D GI products which offer satisfying and effective solutions for the end users.

[2] By employing Likert-type items, there is an assumption that there is an underlying continuous variable within the respondents' attitudes. Likert-type items are inherently ordered categories (e.g. on a scale of 1 to 10) used to indicate the degree of agreement with a statement. However, to analyse this, an interval scale must be used. This raises the question if Likert-type results (ordinal data), converted to numbers, can be treated as interval data. Further, the value assigned to each Likert-type item is arbitrary and dictated by the researcher. It is debateable whether the 'distance' between successive item categories are equal e.g. Is the difference between '2 - Slightly useful' and '3 - Moderately useful' equivalent to the distance between '4 - Very useful' and '5 - Extremely useful'? Despite Likert-type items being technically ordinal, for all intent and purposes, assigning a numerical value and treating it as interval data can provide useful insight not previously possible.

addressed. While the GIS practitioners may have been able to understand the technical terminology and better articulate their needs at the user level, they may not be able to comment about the wider decision-making context. Future work with other participants beyond this group will be beneficial.



3. Please rate the usefulness of the following 3D information according to your day-to-day work:

	Extremely useful	Very useful	Moderately useful	Slightly useful	Not at all useful		Not applicable
Roof geometry	○	○	○	○	○		○
Windows & doors geometry	○	○	○	○	○		○
Texture and/or photo	○	○	○	○	○		○
Interior geometry	○	○	○	○	○		○
3D road geometry	○	○	○	○	○		○
Maximum roof height	○	○	○	○	○		○
Base of roof height	○	○	○	○	○		○
Trees & other biomass geometry	○	○	○	○	○		○
Underground utilities geometry e.g. electricity/telephone cables	○	○	○	○	○		○
Street furniture geometry	○	○	○	○	○		○
Roof shape type (e.g. hipped, mansard, etc)	○	○	○	○	○		○
Number of floors (building)	○	○	○	○	○		○
Ownership and cadastral information	○	○	○	○	○		○
Address with 3D location e.g. identify the floor or height	○	○	○	○	○		○
Landmarks e.g. statues, key buildings	○	○	○	○	○		○
Bridges, flyovers and underpasses	○	○	○	○	○		○

Figure 1. An excerpt of the ‘Usefulness of 3D’ web-based questionnaire.

3.3 Ethics and data protection

Only summarised results are presented in this study, to ensure that individual participants cannot be identified, allowing for open and candid discussions[3].

3.4 Analysing the results

To identify both potential relationships between individual features and 3D GI applications, as well as potential ‘product groups’, several statistical methods were applied. Descriptive analysis (mean and mode) was used to provide a general overview of the state of 3D GI in the UK and by sector, and where respondents selected more than one category, their responses were duplicated for each sector.

Correlation analysis (Kendall’s tau-b) allowed further investigation into any correspondence and congruence within the participants responses and to determine the relationship between the perception of usefulness of 3D features. Kendall’s tau-b is a non-parametric ranking algorithm which does not make any assumptions about the distribution of the data. It also offers the additional benefit of being able to handle many tied ranks. The coefficient can provide an indication on how the perception on the usefulness of two 3D datasets may fluctuate together e.g. if a participant finds roof geometry useful, they may also find roof shape type useful too. This was used to determine the relationship between the usefulness of different 3D features, agnostic of application.

Unsupervised cluster analysis (*k*-means & TwoStep) was then used to determine the minimum number of potential 3D product groups which satisfy the multiple 3D product approach. Cluster analysis or data clustering is a multivariate method that can form classes of objects with similar characteristics, based on a set of measured variables. Three common procedures are TwoStep, Hierarchical and *k*-means, with each employing a different algorithm for creating clusters. Each procedure has its advantages and disadvantages which are described in detail in each section below. From a data type perspective, Hierarchical cluster analysis is limited to small datasets, while *k*-means is restricted to continuous values. TwoStep can create clusters based on both continuous and categorical variables. Further, one of the main benefits is the algorithm’s ability to automatically determine the optimal number of clusters by comparing the values of a model-choice criterion across different solutions. However, only complete cases can be considered. To ameliorate the effects of incomplete cases or missing data, the use of *k*-means allows cases to be retrospectively assigned to a cluster based on distances that are computed from all variables with non-missing values. Nearest-neighbour assignment for partial data was adopted after complete case analysis, resulting in far fewer cases being omitted. To exploit the strengths of different clustering algorithms and to allow for cross comparison, both *k*-means and TwoStep clustering were conducted, and the results compared.

Exploratory factor analysis was used to identify if any latent factors could be used group the variables. Factor analysis is a method to describe variability among observed, correlated variables in terms of a potential lower, latent number of unobserved variables (factors). It works by grouping similar variables into dimensions called ‘factors’. This allows the identification of any potential underlying groups of the 3D features, providing an indication of possible product groups.

Analysis was carried out using Oracle Database 11g for data storage and the IBM SPSS Statistics 24 package for correlation analysis, cluster analysis and exploratory factor analysis.

4. RESULTS

4.1 Summary of the questionnaire results

The questionnaire was conducted in May 2017. A total of 202 completed responses were received out of 532 questionnaire views, representing a 37.67% cooperation rate.

As country-specific information was required (see Section 2.3 and also Wong and Ellul 2017), the country of work was inferred using the supplementary information provided by the participants. Specifically, the participants’ organisation’s name (Q1), sector (Q2), as well as their email address (where supplied) was used. Table 1 shows the spread of the participants’ country of work, with just under 60% from the United Kingdom. Responses were classed as ‘Unknown’ when a country of origin could not be determined or inferred.

Figure 2 shows the distribution of participants by sector. Government and local council (37), Infrastructure and transport

---

[3] This study is compliant with the Data Protection Act 1998. Data Protection Registration Number: Z6364106/2016/01/27. UCL Ethics Project ID Number: 8319/001.

(20) and Academia (19) were the most represented. Sectors falling under the 'Other' category included: Architecture, Charity, Defence, Faith, Healthcare, Marine and, Personal. There were no participants from the UK for Arts and Entertainment, Forestry, or Navigation and routing. This could be due to the lack of GI users from these fields or they could not be reached via the dissemination method. As GIS is interdisciplinary in nature, the questionnaire allowed participants to identify as working in more than one sector. By allowing multiple selections, it was also intended to reduce participants' frustration, thus improve response rates. This design choice, however, led to an issue when sorting the 122 completed UK responses into their respective sectors. Subsequently, where respondents selected more than one category, their responses were duplicated for each sector, resulting in a total of 189 responses.

For the main question, the participants were asked to rate the 'usefulness' of different 3D information from a suggested list, according to their day-to-day work. Figure 3 shows an aggregated summary of the results of UK-only responses. From an initial inspection, ownership and cadastral information (29%), underground utilities geometry (24%) and address with 3D location (24%) are the top three 3D information found to be 'Extremely useful'. Conversely, windows and doors geometry (21%), interior geometry (21%) and texture and/or photo (20%) were described to be 'Not at all useful'. Considering the application-specific and country-dependent nature of 3D user requirements, these are simply initial aggregated descriptors. To fully analyse the results, the responses was split by sector.

Table 1. Participant's inferred country of work

Country	Count	%
UK	121	59.9%
Non-UK	59	29.2%
Unknown	22	10.9%
TOTAL	**202**	

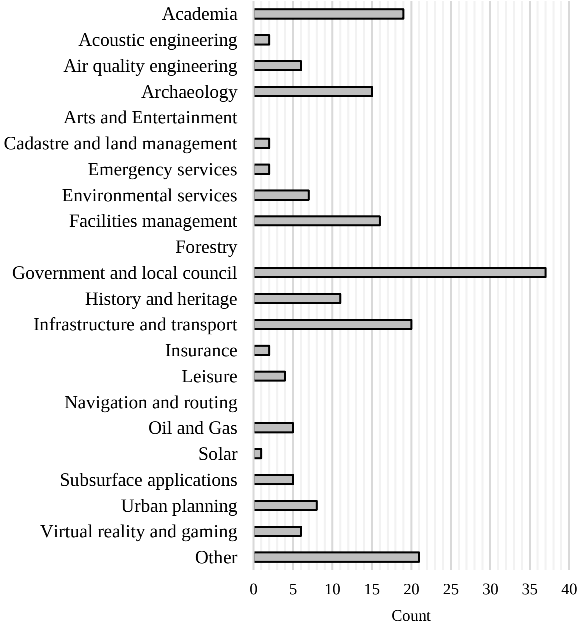


Figure 2. Which sector would you describe yourself to be in? – UK responses

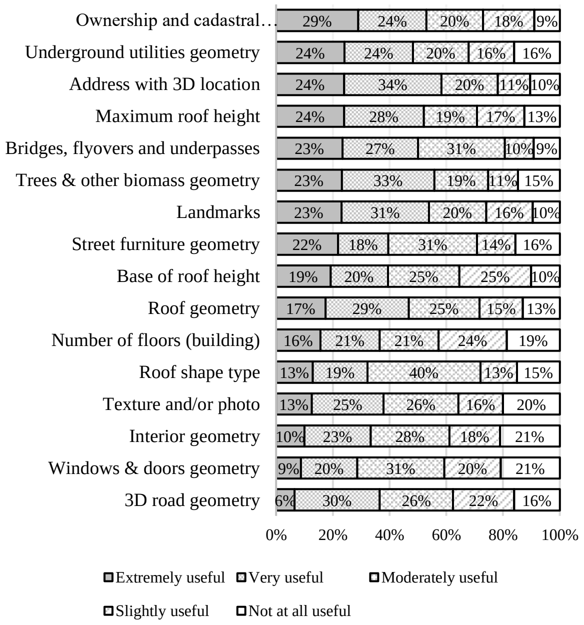


Figure 3. Stacked bar chart showing the aggregated results from question 3 as a percentage split of UK-only responses – 'Please rate the usefulness of the following 3D information according to your day-to-day work?'

4.2 Descriptive analysis

Two measures of central tendency were calculated, the mode and the median. Examining the mode, of the 11 sectors above the n=5 response threshold, 'Air quality engineering', 'Infrastructure & transport' and 'Environmental services' presented the most positive outlook on the usefulness of 3D. Almost all 3D datasets were considered 'Very useful' or 'Extremely useful' with one exception – 'Texture and/or photo' is considered 'Not at all useful' within the context of 'Environmental services'. Examining the less positive side of the spectrum, 'Oil and Gas', 'History and Heritage' and 'Archaeology' presented the least favourable outlooks. Unsurprisingly, above ground information such as 'Roof geometry' and 'Roof shape type' was deemed to be not relevant to the predominantly subsurface work of the 'Oil and Gas' sector. Examining the median is useful for ameliorating the effects of extreme values. Overall, the results were similar to the mode, albeit with a few differences. Notably, for 'Environmental services', 'Windows and doors geometry' (2.5) and 'Interior geometry' (2.5) were viewed as less useful when considering the median versus the mode response (both 5.0).

4.3 Correlation

As noted above, not all 3D features are relevant or 'useful' to all users. Kendall's tau-b was used to examine potential correlation, with the results filtered by a threshold value to identify strongly correlated features. A threshold value of 0.5 was used for Kendall's tau-b, selected by filtering at multiple levels (0.3, 0.4, 0.5, 0.6) and identifying an optimum value in an exploratory manner. Table 2 shows a summary of features and corresponding correlated 3D information, sorted in descending order. For each 3D feature, any other feature with a correlation >0.5 is listed. The interpretation is straightforward – for example, if a user finds 'base of roof height' information to be useful, they have a high chance of also finding 'roof geometry', '3D road geometry' and 'maximum roof height' useful too.

By correlating the responses and filtering at an appropriate threshold, features begin to group together. For example, using Kendall's tau-b, the responses for 'Interior geometry' correlate with those for 'Windows & doors geometry'. It is therefore inferred that should users be interested in detailed indoor information within a mapping product, they would also desire information on the exterior façade and external windows and doors.

Table 2. Summary table of features with a Kendall's tau-b >0.5, sorted in descending order.

3D feature	Correlated featured (>0.5)	Kendall's tau-b
Roof geometry	Base of roof height Roof shape type Maximum roof height Number of floors 3D road geometry	0.69 0.69 0.62 0.53 0.51
Windows & doors geometry	Interior geometry	0.62
Texture and/or photo	-	-
Interior geometry	Windows & doors geometry	0.62
3D road geometry	Bridges, flyovers and underpasses Base of roof height Street furniture geometry Roof geometry	0.55 0.52 0.52 0.51
Maximum roof height	Roof geometry Base of roof height Trees & other biomass geometry Roof shape type Number of floors	0.62 0.74 0.57 0.52 0.65
Base of roof height	Roof geometry 3D road geometry Maximum roof height	0.69 0.52 0.74
Trees & other biomass geometry	Maximum roof height Street furniture geometry Bridges, flyovers and underpasses	0.57 0.57 0.54
Underground utilities geometry	-	-
Street furniture geometry	3D road geometry Trees & other biomass geometry Bridges, flyovers and underpasses	0.52 0.57 0.58
Roof shape type (e.g. hipped, mansard, etc)	Roof geometry Maximum roof height Number of floors (building)	0.69 0.52 0.56
Number of floors (building)	Roof geometry Maximum roof height Roof shape type Address with 3D location	0.53 0.65 0.56 0.50
Ownership and cadastral information	Address with 3D location	0.66
Address with 3D location e.g. identify the floor or height	Number of floors Ownership and cadastral information	0.50 0.66
Landmarks e.g. statues, key buildings	-	-
Bridges, flyovers and underpasses	3D road geometry Trees & other biomass geometry Street furniture geometry	0.55 0.54 0.58

Table 3 shows potential product groups that could be formed from the correlations. The table contributes towards addressing the second sub-question of this study by identifying a set of possible 3D products in a multi-product approach for 3D mapping. Specifically, the table shows five potential product groups which are complementary to each other e.g. a user could combine both basic 3D building information with 3D roads.

Table 3. Potential product groups and features derived from the Kendall's tau-b correlation

Product group	Features
Basic building information	Roof geometry Roof shape type Base of roof height Maximum roof height Number of floors
Detailed building geometry	Windows and doors geometry Interior geometry
Roads	3D road geometry Bridges, flyovers and underpasses Street furniture geometry Trees & other biomass geometry
Land ownership and addressing	Ownership and cadastral information Address with 3D location
Standalone features	Underground utilities geometry Texture and/or photo Landmarks

4.4 Cluster analysis

Part of the aim of this part of the study was to not only to understand the variety of user requirements, but also to begin to classify applications into product groups which demonstrate similar 3D requirements. These groups would, in turn, help inform initial product prototypes. To test this further, cluster analysis was used to assess the groupings of the data. Cluster analysis is a multivariate method that can form classes of objects with similar characteristics, based on a set of measured variables. In this study, TwoStep and *k*-means cluster analysis was used.

For the TwoStep cluster analysis, only 123 fully completed UK responses were used within the clustering. Multiple iterations were performed, using different variables and parameters. The first iteration resulted two clusters, with a silhouette measure of cohesion and separation of 0.450 indicating fair cluster quality. However, upon inspecting the predictor importance, roof-based information (e.g. maximum roof height, roof geometry) was found to dominant as the main variables in estimating the model. Therefore, for subsequent iterations, roof-based characteristics were aggregated to ameliorate this bias. Table 4 shows the results of the TwoStep clustering. The table shows the changes in cluster composition and goodness-of-fit (silhouette measure of cohesion and separation or silhouette coefficient) in successive clustering iterations. The silhouette measure of cohesion and separation is a measure from 0 to 1, showing cluster quality. A value tending to 1 indicates a good quality cluster with small within-cluster distances and large between-cluster distances. The 'ratio of sizes' indicate the difference in size between the largest and smallest cluster. Ideally, clusters are of equal sizes (Milligan et al., 1983). A rule of thumb proposed is that no single cluster is twice as large as any other cluster. The results from Table 4 show that both ratio of sizes and cluster quality decreases with increasing cluster size. The results from the TwoStep cluster analysis indicate that two clusters offer the optimal solution. In addition, a *k*-means clustering with two, three, four and five clusters with a maximum of 10 iterations. Roof-related variables were aggregated (as per the TwoStep clustering). Convergence was reached for all within the threshold except for the four-cluster solution. The *k*-means results showed that, like the TwoStep clustering, two clusters yielded the most satisfactory solution. The silhouette coefficient (0.307), however, only showed a weak clustering structure. The two-cluster solution was investigated further, by reassigning the responses into their respective cluster membership. The results (for both TwoStep and *k*-means) showed that there is a dichotomy within almost all the sectors. This shows that appetite for 3D information not only varies between sectors, but also within sectors between different practitioners.

Table 4. Change in cluster size and quality over different cluster variations for TwoStep and *k*-means.

Iteration	Clusters	Ratio of sizes	Silhouette measure of cohesion and separation
		TwoStep	
TS-1	2	1.95	0.44
TS-2	3	1.43	0.33
TS-3	4	2.05	0.35
TS-4	5	2.44	0.31
		***k*-means**	
KM-1	2	1.54	0.31
KM-2	3	6.56	0.21
KM-3	4	9.55	0.21
KM-4	5	8.75	0.15

4.5 Exploratory factor analysis

Data were subject to factor analysis using Principal Axis Factoring and orthogonal Varimax rotation. The Kaiser-Meyer-Olkin measure (KMO) of sampling adequacy was 0.86 was well above the commonly recommended value of 0.5 indicating that the data were sufficient for EFA. The Bartlett's test of sphericity χ^2 was significant (120) = 2256.69, $p < 0.001$ showed that there were patterned relationships between the items. Using an eigenvalue cut-off of 1.0, there were three factors that explained a cumulative variance of 58.493%. The scree plot confirmed the findings of retaining three factors. The table below shows the factor loading after rotation using a significant factor criterion of 0.4 (Table 5). The three factors were given the names 'Non-building information', 'Detail building information' and 'Simple building information'.

Table 5. Truncated summary table of the EFA.

	Simple building info.	Detailed building info.	Non-building info.
Underground utilities geometry			0.76
Street furniture geometry			0.76
Bridges, flyovers & underpasses			0.67
Trees & other biomass geometry			0.66
Ownership & cadastral information			0.57
Address with 3D location			0.55
Landmarks			0.54
3D road geometry			0.50
Windows & doors geometry		0.76	
Interior geometry		0.67	
Texture and/or photo		0.63	
Roof shape type	0.54	0.59	
Roof geometry	0.55	0.55	
Base of roof height	0.78		
Number of floors	0.64	0.49	
Maximum roof height	0.64		0.41
Eigenvalues	**2.77**	**2.88**	**3.71**

Table 6. Median response for UK participants split by factor and sector, sorted by the sum.

Sector	n	Simple building info.	Detailed building info.	Non-building info.	Sum
Solar	1	4.3	4.4	4.8	13.5
Air quality eng.	6	4.5	3.6	4.4	12.5
Subsurface apps.	5	3.7	3.5	4.5	11.7
Cad. & land mgmt.	2	3.8	3.7	4.1	11.6
Acoustic engineering	2	4.0	3.7	3.6	11.3
Env. services	7	4.0	3.2	4.0	11.2
Facilities mgmt..	16	3.5	3.8	3.9	11.2
Urban planning	8	4.0	3.2	3.9	11.1
VR & gaming	6	4.3	3.1	3.5	10.9
Infra. & transport	20	3.3	3.3	4.1	10.7
Other	21	4.0	3.0	3.6	10.6
Academia	19	3.8	3.2	3.4	10.4
Archaeology	15	3.3	3.5	3.3	10.1
Gov. & local council	37	3.3	2.9	3.6	9.8
History & heritage	11	3.0	3.4	3.1	9.5
Emergency services	2	2.3	2.4	2.5	7.2
Oil & gas	5	2.5	1.8	2.9	7.2
Leisure	4	2.0	1.8	3.1	6.9
Insurance	2	1.7	1.7	1.4	4.8

5. DISCUSSION

5.1 Features vs. applications

As an overview, the mode and median responses showed that GI users from infrastructure and transport, air quality engineering and environmental service could be potential early adopters of 3D GI. Despite this, the cluster membership result show that even within sectors, there is a split between participants who perceive 3D to be moderately to extremely useful, and others who perceive it to only be slightly to moderately useful. The findings from this study, however, could not demonstrate a current need to produce a 3D dataset with complex and detailed building geometry. Rather, simple buildings coupled with non-building classes are desired by the user. It is important to bear in mind that this does not mean enhancements to a building's geometry (such as the position of windows and doors and the roof shape), are not important, but rather it is not a current priority. Instead, it is suggested that the current drive should be towards simple buildings coupled with non-building classes (e.g. vegetation and street furniture) and building-based attribution. The finding from this study is in contrast to the current trend in within academia and data producers in acquiring more detailed roof geometry and LoD2 representations. However, it is important to stress that practitioners may currently value non-spatial data items as they are easier to use and incorporate within their workflow. Detailed 3D geometric information is still challenging to manage and exploit, and may therefore be considered as less useful for now. In time, this may change, with improved software, hardware and processes to handle 3D GI.

Regardless of the application, the results showed that participants perceived non-building classes to be more useful than additional detail on building geometry. For example, inspecting the median response values for Urban Planning participants showed that building-related enhancements such as 'Windows and doors geometry, 'Texture and/or photo' or 'Interior geometry' were less desired than other non-building thematic classes including 'Tree and other biomass', 'Underground utilities geometry', 'Street furniture geometry', 'Landmarks', 'Address with 3D location' and 'Bridges, flyovers and underpasses'. Some basic building-based classes were still desired however, such as 'Roof geometry', 'Maximum roof height' and 'Roof shape type'. Exploratory factor analysis further supports this finding as three distinct groups emerged: 1) Simple building information; 2) Detailed building information; and 3) Non-building information. Relating these groups back to the applications and creating composite median scores further supports the idea that simple building geometry coupled with non-building thematic classes is perceived to be most useful for users. This finding reflects the expectation from Biljecki et al. (2015) that more use cases will take advantage of thematic features other than buildings in the future. Despite the perception of non-building features to be useful, there are currently very few guidelines on features such as roads and street space as the focus has been on modelling buildings (Beil and Kolbe, 2017). On-going work on modelling vegetation (e.g. trees and root systems) are being conducted (Iñiguez, 2017). Further work is therefore required to establish standards for modelling non-building features in 3D.

5.2 One-size-fits-all vs. multiple product approach

Cluster analysis was conducted to group applications with similar requirements, on the assumption that multiple 3D mapping products may be more suitable than a single mapping product. However, both the TwoStep and *k*-means clustering resulted with an optimal two-cluster solution. On one hand, this may indicate there are no clear product groups. Another interpretation is that it reflects the uncertainty within the users on what 3D information is useful for their application. The lack of clarity could alternatively be attributed to the use of 5-point Likert items. Despite literature supporting the use of 5-point Likert items to increase response rate, response quality and reduce respondents' 'frustration level', the use of a 7-point or even 9-point scale could provide more points of discrimination. Conversely, having seven or nine points may lead to ambiguous responses as even in the case of a five-point scale, the distinction between 'extremely useful' and 'very useful' is not always immediately clear or consistent between participants. Despite the cluster analysis not offering a clear result, the correlation analysis was able to provide five initial groups of 3D features (see Table 3). These offer a starting point to further validate whether a modular, multi-product approach to 3D GI production is most efficient.

5.3 Visualisation vs. 3D data and analysis

It is important to distinguish between the requirements related to visualisation and requirements related to 3D data. For example, to support visualisation-based decision-making processes, a geometric mesh with photorealistic texture mapping may be sufficient. Conversely, texturing is less important within 3D analysis. Therefore, simple untextured, but structured 3D geometry at a wider coverage may be required. This difference in data requirements can be problematic for NMCAs and other producers as they required different acquisition methods and data structures. One potential solution could be to produce two different products – the City of Helsinki produces both a semantic city information model (for analyses and simulation) and a 'reality mesh' or photogrammetric mesh (for visualisation). By having both datasets, the city can hedge their bets and draw the strengths from both representations, thus future proofing their product line.

5.4 Requirements gathering challenges

Part of the research approach was to ensure as many potential users of 3D GI was covered. Within the sample, there was some overrepresentation in certain sectors and underrepresentation in others. In sectors which were underrepresented, it was not possible to conclude any findings as a minimum threshold of five responses was used. To compound the issue, the interdisciplinary nature of GIS means that participants often work in more than one sector. It was therefore possible to identify as working in multiple applications within the questionnaire. Forcing a user to identify from a single sector, however, would have increased user frustration, and inevitably, survey abandonment. This compromise, however, meant that some responses had to be repeated when dividing by sector e.g. if a participant identified as both Archaeology and Academia, their responses were duplicated for each sector. Despite the inherent extra weighting given to certain responses, only 38 of 121 UK responses identified itself as more than one sector, and only 16 identified as three or more sectors.

5.5 Analysis of research approach

Within this study, a number of exploratory statistical analysis techniques were used to identify both potential relationships between individual features and 3D GI applications, and potential 3D product groups. Descriptive and correlation statistics were effective in distilling clear trends from the results of the questionnaire, providing a summative overview of the state of the use of 3D in the UK. Leveraging the results from the correlation analysis also allowed potential product groups to be formed, by interpreting the responses agnostic of the user's application. Unsupervised clustering was intended to provide a similar function, although the quality of the clustering results was not entirely satisfactory. Further work is required, such as with larger sample sizes, to validate if clustering is an appropriate methodology for eliciting user requirements for 3D GI.

5.6 Practical implications for Ordnance Survey and other NMCAs

As described by Capstick et al. (2007), 'The first step in the creation of a 3D geospatial infrastructure is the definition of the data specification.'. In order to define a data specification, Ordnance Survey must understand the broad range of end-users of 3D data as well as how economic and feasible it is to collect such data. While the end-user requirements aspect is, in part, addressed in this study, the economic and 'feasibility' aspect of product such data is not fully investigated. From a production perspective, even creating simple 3D models, is not inexpensive nor trivial, as it still requires a large amount of manual effort (Sargent et al., 2015). In addition, there is still little guidance and research on how to capture, model, and structure non-building features in 3D. There is therefore a need to understand the actual costs of production of different types of 3D features and enhancements in the context of Ordnance Survey. In particular, there is a need to compare between what features and enhancements are easy (and cheap) to produce, and what are actually useful and desired by the end-user (as identified in this study). For example, as many NMCAs already capture data from airborne platforms in nadir view for 2D mapping, the production of roofs and roof-based features such as chimneys are comparatively easier than to capture windows and doors geometry (which may require a mobile mapping system).

It is important to re-emphasise here that uptake of 3D applications varies from country-to-country and that different nations are bound by different organisational and structural limitations. structures. Results from this study are in the context of Ordnance Survey (and Great Britain), and may therefore not be generalisable to other countries. Other national mapping agencies may not have the same unique characteristics of Ordnance Survey, nor do their nations have the same political landscape. Where NMCAs have already conducted their own surveys on customer needs (Lantmateriet, Sweden and GUGiK, Poland), the findings from this study can provide a point of comparison. Where NMCAs have yet to conduct any requirements gathering exercise, the methods and approaches used in this study can be easily replicated and adapted for their country. This will allow other national mapping agencies who are looking to establish their own 3D national mapping product (or products) to capture country-specific 3D requirements.

5.7 Further work

The work presented in this study presented the first iterations of user requirements gathering for 3D GI. Several opportunities for future work arose from this study. Firstly, repeating the exercises with a larger sample and with non-GI users would be beneficial. The challenge here is in capturing a representative sample of adequate size to encompass the multitude of GI-applications as well as acquiring enough detail to elicit detailed requirements. It would also provide a validation for the requirements collect so far. Secondly, repeating the exercises in 1, 5 and 10 years' time would be valuable in assessing any change in requirements. As technology evolves and improves, what were previously barriers may no longer exist, and as 3D data becomes more common place the end users' understanding of their requirements will evolve – i.e. it is important to remain up-to-date with the requirements of the user. Lastly, other requirement elicitation methods such as on-site observation and focus groups may be beneficial in later iterations of the development life cycle, allowing for a deeper understanding of the user. Creating and testing prototypes will allow the confirmation and refinement of the user requirements over time. Note that this must not be a linear process; for 3D GI, new use cases are expected to emerge over time thus there is also a need for continued requirements gathering to capture any new applications.

6. CONCLUSION

Understanding the user perception of 'usefulness' is an important step towards designing 3D data that is effective and usable. This study has shown that 'Air quality engineering', 'Infrastructure & transport' and 'Environmental services' present the most positive outlook on the usefulness of 3D GI in the United Kingdom. The results further showed that users perceived non-building classes and building attribution to be more useful than additional detail on building geometry. Cluster analysis showed that appetite for 3D information not only varies between sectors, but also within sectors between different practitioners. While an initial set of five potential product groups and features was derived from correlation analysis, the cluster analysis was unable to offer a clear result of whether a 'one-size-fits-all' or multi-product approach is most suitable for 3D GI production. The current study has only examined existing GI practitioners. Further work is required to incorporate both non-GIS users and 3D users outside the GIS domain.

REFERENCES

Beil, C., Kolbe, T., 2017. CityGML and the streets of New York - A proposal for detailed street space modelling. ISPRS Annals of Photogrammetry, Remote Sensing & Spatial Information Sciences 4.

Biljecki, F., Stoter, J., Ledoux, H., Zlatanova, S., Çöltekin, A., 2015. Applications of 3D City Models: State of the Art Review. ISPRS International Journal of Geo-Information 4, 2842-2889.

Finnish Forest Association, 2016. Finnish forests resources. https://www.smy.fi/en/forest-fi/forest-facts/finnish-forests-resources/.

Forestry Commission, 2017. Forestry Statistics and Forestry Facts & Figures. https://www.forestry.gov.uk/forestry/infd-7aqdgc.

Iñiguez, J.G., 2017. Geometric Modelling of Tree Roots with Different Levels of Detail. ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences 4, 29.

Milligan, G.W., Soon, S.C., Sokol, L.M., 1983. The effect of cluster size, dimensionality, and the number of clusters on recovery of true cluster structure. IEEE transactions on pattern analysis and machine intelligence, 40-47.

Sargent, I., Holland, D., Harding, J., 2015. The building blocks of user-focused 3D city models. ISPRS International Journal of Geo-Information 4, 2890-2904.

Stoter, J., Streilein, A., Pla, M., Baella, B., Capstick, D., Home, R., Roensdorf, C., Lagrange, J., 2013. Approaches of national 3D mapping: Research results and standardisation in practice, 8th 3DGeoInfo Conference & WG II/2 Workshop, Istanbul, Turkey, 27–29 November 2013, ISPRS Archives Volume II-2/W1. ISPRS.

Stoter, J., Vallet, B., Lithen, T., Pla, M., Wozniak, P., Kellenberger, T., Streilein, A., Ilves, R., Ledoux, H., 2016. State-of-the-art of 3D National Mapping in 2016. Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci. XLI-B4, 653-660.

Tang, P., Huber, D., Akinci, B., Lipman, R., Lytle, A., 2010. Automatic reconstruction of as-built building information models from laser-scanned point clouds: A review of related techniques. Automation in construction 19, 829-843.

Tuokko, J., 2017. Value Chain Mapping – Forestry Management Use Case, the Added Value of 3D Geo-information, EuroSDR workshop on "Identifying the economic value of 3D geoinformation", 30th – 31st March 2017, Barcelona.

Walter, V., 2014. A survey on state of the art of 3D Geographical Information Systems. European Spatial Data Research, EuroSDR Officla Publication No. 64, 66-113.

Witmer, R., 2017. Value Chain Mapping – 3D Cadastre and Valuation, the Added Value of 3D Geo-information, EuroSDR workshop on "Identifying the economic value of 3D geoinformation", 30th – 31st March 2017, Barcelona.

Wong, K., Ellul, C., 2017. User Requirements Gathering For 3D Geographic Information In The United Kingdom. ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences 4, 125.