This is an Accepted Manuscript of an article published by Taylor & Francis in the *International Journal of Geographical Information Science* in 2022, available online: https://doi.org/10.1080/13658816.2022.2140808. Cite as:
Lei B, Stouffs R, Biljecki F (2022): Assessing and benchmarking 3D city models. *International Journal of Geographical Information Science.*
It is deposited under the terms of the Creative Commons Attribution-NonCommercial License (http://creativecommons.org/licenses/by-nc/4.0/), which permits non-commercial re-use, distribution, and reproduction in any medium, provided the original work is properly cited.

Assessing and benchmarking 3D city models

Binyu Lei[1], Rudi Stouffs[1], and Filip Biljecki[1,2,*]

[1] *Department of Architecture, National University of Singapore, Singapore — binyul@u.nus.edu; stouffs@nus.edu.sg; filip@nus.edu.sg*
[2] *Department of Real Estate, National University of Singapore, Singapore*
[*] *Correspondence at Department of Architecture, National University of Singapore, 4 Architecture Dr, 117566 Singapore, filip@nus.edu.sg*

Abstract

3D city models are omnipresent in urban management and simulations. However, instruments for their evaluation have been limited. Furthermore, current instances are scattered worldwide and developed independently, hampering their comparison and understanding practices. While there are developed assessment frameworks in open data, such efforts are generic and not applied to geospatial data. We establish a holistic and comprehensive 4-category framework '3D City Index', encompassing 47 criteria to identify key properties of 3D city models, enabling their assessment and benchmarking, and suggesting usability. We evaluate 40 authoritative 3D city models and derive quantitative and qualitative insights. The framework implementation enables a comprehensive and structured understanding of the landscape of semantic 3D geospatial data, as well as doubles as an evaluated collection of open 3D city models. For example, datasets differ substantially in their characteristics, having heterogeneous properties influenced by their different purposes. There are further applications of this first endeavour to standardise the characterisation of 3D data: monitoring developments and trends in 3D city modelling, and enabling researchers and practitioners to find the most appropriate datasets for their needs. The work is designed to measure datasets continuously and can also be applied to other instances in spatial data infrastructures.

Keywords: 3D GIS, open data, digital twins, open government, data quality

1 Introduction

3D city models have increasingly developed in the past years. Numerous studies have affirmed their advantage in supporting different stakeholders, e.g. urban planners, citizens, and researchers. As the technologies and supporting infrastructure have been strengthening, many cities and countries adopted 3D city models as an advanced tool to solve urban issues, catalyse the process of participation, and inform decision-making across a variety of use cases (Arroyo Ohori, 2020; Biljecki et al., 2021; Dukai et al., 2020; Gil, 2020; Labetski et al., 2022; Ledoux et al., 2021; Liang et al., 2020; Noardo et al., 2022;

Palliwal et al., 2021; Park et al., 2021; Stoter et al., 2020; Virtanen et al., 2021; Vitalis et al., 2020; Wysocki et al., 2021). Subsequently, over the past few years, there has been a noticeable increase in the number and variety of 3D city models publicly released by different organisations, primarily governments.

With the involvement of different stakeholders, geographical coverage, and proliferation of an increasing breadth of use cases, 3D city models are scattered around the world and developed in disconnected initiatives, which leads to heterogeneous characteristics (Bognár et al., 2021). Currently, governments are the major provider of 3D city models, but efforts by universities and research institutions also present a fair share in contributing to the availability of datasets, while crowdsourcing has been gaining currency as well (Wu & Biljecki, 2022). This diversity entails that 3D city models nowadays vary in methods of data acquisition, processing, storage, publication, use, and maintenance along with technical, socio-economic, political and cultural differences (Conradie & Choenni, 2012; Dukai et al., 2020; Peters et al., 2022). These issues make it difficult to understand practices, characterise a *good* 3D city model, gauge whether it is suitable for a particular use case, and explore the integration of multiple datasets. It may also lead to ambiguity in comparing multiple datasets, e.g. given a choice between two datasets, as it is not clear which one is *better*. This topic is also relevant in the context of urban digital twins, for which there is a consensus that 3D city models are a pillar (Shahat et al., 2021).

At the moment, there is no holistic and standardised instrument to understand the characteristics of 3D city models, which would bridge the aforementioned issues. Despite some recent efforts focused on particular aspects such as integrity of geometry and topology (such related work will be overviewed in Section 2.1), there has been no mechanism to characterise 3D city models by their properties that hint at their quality, usability, and suitability.

Take the language barrier as an example. Although the documentation and content of many 3D city models are available in English, many still use local languages, which may hinder understanding and interoperability (e.g. mismatch of attribute names and values, which may not be recognised by software). While plenty of use cases, practitioners and researchers rely on high-quality open data, many public administrations do not provide integrated and standardised 3D city models, which makes further analysis difficult in practice (Susha et al., 2015). 3D city models are often managed in different standards with different sets of information in the datasets to support various purposes. Moreover, while many applications of 3D city models may have some common requirements with each other, datasets may be heterogeneous and not all of them are suitable for particular use cases. An issue is compounded by the lack of means to characterise data and their fit for purpose. For example, some visualisation-only applications do not require semantics, while others demand particular attributes to conduct analyses (Lim et al., 2020). Among various implications, it is challenging for stakeholders to understand the best practices and determine the characteristics of a dataset they seek to acquire. As various purposes have shaped the provenances of different datasets, there is no general manner nor specific standard to produce 3D city models, leading to diverging datasets.

Apart from the heterogeneous nature of 3D city models hampering comparison, a lack of a standardised approach to characterise datasets is another issue. Although the research fields of, e.g. gauging open government data, have been well established (Kedron et al., 2021), such research does not apply to 3D geoinformation. In terms of open data released by the authorities, many studies evaluate primarily three main aspects: quality, adoption, and barrier (Borzacchiello & Craglia, 2012; Conradie & Choenni, 2012; Degbelo, 2021; Susha et al., 2015). There are several generic initiatives for understanding the characteristics and usability of open data instances, (e.g. Global Open Data Index[1], Open Data Barometer[2]), and some research papers (Braunschweig et al., 2012; Strathern, 2000) provide insight into various aspects, such as the quality. However, in terms of 3D data in the geospatial realm, there lacks a comprehensive approach to realise a related but specific implementation.

The motivation of this study is to build a universal framework to measure relevant properties of 3D city models, which can assist in gauging their usability, quality, and development. Further, in an aggregated manner, such insights can provide an understanding of the trends of capturing, maintaining, and releasing 3D geoinformation. The framework we design is intended to apply to instances of 3D data that are either closed or open. The same goes for their lineage, applying to either governmental or non-governmental data, and the different ways to acquire them. However, in the implementation of the framework in this paper, only open government data is taken into account to ease obtaining data. To our best knowledge, this framework is the first effort to characterise 3D city models and compare their properties. Furthermore, such a framework will lead to formalising the suitability of data for a particular use case (e.g. developers of a simulation software may use our framework to specify minimum requirements for a dataset for a reliable analysis). Other applications are many. For example, the framework would facilitate identifying particular characteristics of a dataset that may be useful to researchers and users for whom the geography of the dataset is not important. They may find a dataset with a particular characteristic important for their development regardless of its location (e.g. developers of a wind flow simulation software that are looking for 3D data around the world with certain requirements to test the development). As similar work has not been done for other types of geographic information, we believe that it also contributes broadly to the domain of spatial data infrastructures and may be adapted to suit other instances of spatial information.

We address three objectives: (1) define relevant properties of 3D city models; (2) establish a framework to characterise 3D city models; and (3) benchmark and compare the performance of a selected set of datasets as the implementation of the framework. The last point also doubles as an instrument to gauge the state of the art of 3D geoinformation and to provide a collection of 3D city models around the world together with their properties.

[1] https://index.okfn.org

[2] https://opendatabarometer.org/?_year=2017&indicator=ODB

2 Background and related work

2.1 Related work in 3D GIS assessment

While there is a substantial gap in characterising and evaluating 3D city models, there has been working on validating particular aspects of 3D datasets, primarily the geometric integrity (Alam et al., 2014; Dukai et al., 2021; Ennafii et al., 2019; Ledoux, 2018; Wagner et al., 2015) and adherence to standards such as CityJSON and CityGML (Biljecki et al., 2016; Ledoux, Ohori, et al., 2019). Geometric and schematic consistency is not in our domain, as our approach is geared towards a broad understanding of the characteristics of 3D city models. However, such work is relevant to us as it may complement ours (geometric quality is a consideration that may affect usability). Some researchers note that 3D city models have application-specific requirements (Biljecki & Tauscher, 2019; Coors et al., 2020; Tang et al., 2020), which we regard in our work to develop a framework that will aid researchers in gauging the usability of a dataset.

Another relevant work is the one by Labetski et al. (2018), developing means to describe metadata of 3D city models. Considering that we aim to identify relevant properties of 3D city models, our work intersects with metadata, so we adopt certain elements from their work that characterise 3D city models, e.g. Level of Detail (LoD), textures, and thematic modules present in a dataset.

2.2 Related work on benchmarking data

Existing approaches to describing and assessing data mostly focus on evaluating open government data efforts (Shannon & Walker, 2018), and they have different emphases depending on particular purposes, with some being general. Berners-Lee (2006) proposed a Five-Star Scheme to assess the usability of datasets, based on (1) availability of a website with any formats openly licensed; (2) availability of being as structured data (e.g. spreadsheet rather than images); (3) availability of non-proprietary formats; (4) use of open standards; and (5) availability of linking with other datasets. Several disciplines have widely adopted the proposed scheme (Hossain et al., 2016; Wilson et al., 2021); however, the scheme mostly concentrates on specific aspects without being holistic and considering other critical properties, e.g. timeliness. Consequently, an outdated dataset may still get a higher score according to their approach.

The work of Kundra (2011) identified ten principles for improving federal transparency when generating and publishing open data. For example, the work states the importance of common data standards, real-time information and collaboration with third parties. Ubaldi (2013) proposed a comprehensive framework to measure open government data with a range of dimensions, covering political, technical, economic, and organisational aspects. In terms of data, the work mainly investigates its availability, quality, uptake and re-usability. In terms of data availability, a number of datasets are assessed whether they are accessible or downloadable on the portals. At the same time, data quality is considered as data accuracy, consistency, frequency of update and affordability. This analytical framework is a valuable attempt to build a broad approach with a large set of metrics from different dimensions. However, it ignores the properties of datasets as all indexes are created at a portal level. Furthermore, a lack of a quantitative standard for assessment makes the work prone to ambiguity.

Based on previous work, Vetrò et al. (2016) set up a framework to assess open government data focusing on data quality. They pick up a set of metrics: traceability (e.g. metadata availability in terms of creation/update process), currentness (e.g. update timeliness), expiration (e.g. delay between creation and publication), completeness (e.g. dataset with complete value), compliance (e.g. dataset with standard), understandability (e.g. comprehensive formats) and accuracy (e.g. accuracy in aggregation and correct value in each cell). For quantitative analysis, they adopt the two-tailed Mann-Whitney test (Sachs, 2012) to compare the results of two examples. This work is a detailed framework to investigate data and set metrics, despite its particular scope of data quality.

Based on such efforts, there have been projects to assess and rank countries or cities in terms of their performance regarding open authoritative data. Since 2013, the World Wide Web Foundation, with the support of the Omidyar Network, has launched the project of Open Data Barometer[3], aiming to reveal global trends of open data initiatives and provide a systematic comparison covering 30 governments. It ranks countries with a tripartite structure regarding readiness, implementation and impacts on government, civil society and business. The three tiers are weighted equally (1/3) in quantitative analysis, with a series of questions to score each sub-index. It is probably the first comprehensive work to evaluate how governments publish and use open data to promote accountability, innovation and social impact. Nevertheless, it remains at quite a high level to assess external performance rather than examine details of the dataset as it emphasises the yearly change of score trends between countries. Another relevant project is the Global Open Data Index[4], conducted by the Open Knowledge Network. This work aims to provide feedback on open government data to gain insight into data gaps and motivate more engagement to make data useful and impactful. The framework is built with a range of categories (e.g. budget, maps, locations and law), and selects a set of questions to examine the openness of each dataset in each city around the world. The scoring system gives 60 points to data accessibility and 40 points to questions around license/status/standard/formats through a survey. As this project intends to measure how datasets are openly published, the final score and the ranking suggest a degree of openness which means only scoring 100% indicates the data is open compared to 80% presenting public but not open.

Certain disciplines have related research, e.g. there is work on assessing data openness in air quality measurements (Mak & Lam, 2021). While these efforts affirm the importance of implementing such work in the domain of GIS or specifically 3D city modelling, they are usually too specific to be transported to our realm. Also, current frameworks in the field of open data maintain two gaps: existing studies lack a comprehensive consideration of data properties, and many have no metrics system to compare datasets in a distinct manner.

[3] https://opendatabarometer.org/?_year=2017&indicator=ODB
[4] https://index.okfn.org

Nevertheless, we take inspiration from the above work — as they cater to general data, and some elements can be adopted in the realm of 3D GIS. To address the limitations described above, we develop a holistic multi-level approach with regard to both portals and datasets, and establish a scoring system to benchmark their performance from a neutral, non-profit, and unbiased academic standpoint.

3 Methodology

3.1 Framework design and principles

We develop a 4-category framework, dubbed '3D City Index' (Table 1) encompassing 47 criteria to gauge 3D models and benchmark intrinsic content (Farhang et al., 2008). We examined research papers and initiatives related to data assessment, screened several evaluation indexes that align with our research objectives, and finally selected these 47 relevant criteria. Such selection aims to reflect the characteristics of 3D city models and help stakeholders understand current models from different aspects. This framework is general and can be applied to different kinds of 3D city models worldwide, e.g. open or closed, governmental or non-governmental, obtained by extrusion or with laser scanning. The implementation of this framework may benefit researchers and practitioners to understand the properties of a dataset, gather the trend of 3D geoinformation, and enable them to choose one or multiple datasets for different purposes. Furthermore, the implementation of framework is intended to drive the future creation of 3D city models. For example, when contracting the acquisition or production of data from companies, governments may set certain requirements using criteria in this framework and enhance internal workflows, i.e. improving openness, increasing the frequency of update, and promoting maintenance services. At the same time, the framework's design is flexible, allowing extensions and customisations for particular scenarios, e.g. use cases.

The proposed framework includes 4 categories and provides a comprehensive understanding to stakeholders spanning a variety of domains and geographies. The structure of this framework starts with an overview of the data portal (Category 1), followed by a description of the 3D city model (Category 2), and then considering its thematic content (Category 3), i.e. object types, and finally examining building attributes in detail (Category 4). The design of our framework includes a high-level description of 3D city models, as well as considering their properties that are relevant for stakeholders and align well with research objectives. With these 4 categories, practitioners may assess datasets from different perspectives, from the portal level to detailed and low-level characteristics.

The framework is data agnostic – we consider 3D geospatial data in all relevant formats. After defining the 4 categories and the scope of this study, we endeavour to gather relevant criteria and obtain 47 criteria under all measurement categories. Each criterion is characterised by a question that has a boolean answer, with 1 point for a yes and 0 for a no. A 'yes' answer indicates that the city model has satisfied this criterion. For example, we propose a criterion for affordability – Is a 3D city model free of charge (2C4). A positive answer here implies that the dataset can be downloaded at no cost. On the other hand, it also indicates that this dataset performs better in openness and accessibility, suggesting that a single criterion may have multiple functions. This quantitative benchmark is intended to serve as a generic approach to assessing 3D city models. Moreover, the framework may also satisfy other purposes for specific uses. That is, it enables giving more weight to a particular aspect that may be more important than another in a particular context. Also, if necessary, the framework may be extended with further aspects, and aspects less relevant for a particular purpose can be omitted.

3.2 Criteria and scoring system

Category 1 is about the data portal and its openness, which is the first step in providing communication and access to a 3D city model. Typically, portals describe datasets, sometimes including visualisations, before their download. The first few criteria consider web-based visualisation and the principles of openness. Web-based visualisation includes a detailed website (1C1), a web viewer (1C2) and the timeliness of provided information (1C3). For example, practitioners could view the 3D virtual city online, search for specific buildings by interest, and obtain further information, such as address and building type. In terms of openness of datasets, it regards the use of language (1C4, 1C5), as well as a way to leave feedback on recommendations or reporting errors (1C6). All criteria are built based on the first phase of facilitating obtaining 3D city models.

Category 2 encapsulates key characteristics of 3D city models. It investigates the availability of datasets, including downloadability, age, frequency of update and the record of historical data, which are highly related to the needs of users (Mak & Lam, 2021). Considering the impact of data acquisition technique, we firstly set two criteria to define city models — whether it is a structured 3D city (information) model (2C1) or a mesh model (2C2), or if it is both. By means of downloadability, the provision of downloads is essential access for subsequent investigation (2C3). Furthermore, an ideally downloadable dataset should be free of charge (2C4) in consideration of access. Downloading data without mandatory registration or additional obstacles is another major consideration (2C5). Requiring additional software for download or approval from data providers hinders the process of procurement (Attard et al., 2015). A dataset provided in more than one format will be more useful in satisfying different purposes (2C6). For example, a dataset in the OBJ format may be better for visualisation than a CityGML, which on the other hand, may be more appropriate for semantic analysis. Thus, instead of favouring a specific format, we attach importance to the availability of multiple formats. Continuing the consideration of data formats, we also take into account whether the dataset has been released according to an open standard (2C7). One of the key benefits of using open data standards is to make city models interoperable, potentially leading to supporting multiple applications (Kolbe & Donaubauer, 2021). To address the practical value, further we include open license (2C8) and metadata (2C9) as additional criteria. The timeliness is highlighted by the criteria of age and update frequency, where it is assumed that recently published datasets potentially represent the built environment more veraciously, and may also be of better quality and accuracy thanks to the continuous advancement in 3D data acquisition. We firstly establish a criterion to assess whether a city model is recently published (2C10), to reflect the age of datasets. The threshold of considering a dataset recent is set to the last 5 years. As it is the case with other criteria, practitioners may customise such details if needed. In terms of update, we consider the status as whether it has been updated at least once (2C11) and define the frequency of update according to the officially stated plans (2C12). These criteria are instrumental as a regularly updated dataset tends to provide accurate information and ensure quality. In addition, we also add one more criterion to describe the version management of datasets: keeping historical datasets (2C13). The availability of historical records may benefit users in studying the development of cities through time. Level of detail (LoD) is another integral component of this category — it indicates the correspondence of a model with its real-world counterpart (Tang et al., 2020). As such, we consider whether the 3D city model includes more than one level of detail (2C14). We also include a criterion of coverage in the framework, as whether it covers the entire jurisdiction (2C15).

In addition to the general features of datasets, we further consider the types of thematic objects in Category 3, which comprehensively reflect physical components of a city. We adopt the Open Geospatial Consortium (OGC) standard CityGML in this category to determine relevant criteria. CityGML is considered as a comprehensive format for representing 3D urban information (Gröger et al., 2012; Kolbe, 2009), and we adopt its taxonomy of thematic features and building attributes, which will be explained in the Category 4. The urban objects regarded by Category 3 are derived from the specification of CityGML. For example, the first criterion (3C1) refers to whether a model is generated with semantically differentiated surfaces. A LoD2 model that has roofs distinguishable from walls is one of the cases. Combining different object types is common in visualisation, planning, management and decision making (Julin et al., 2018), e.g. a dataset including waterbody may enable flood simulations (Schröter et al., 2018).

CityGML also enables representations of the properties of individual urban objects. It is possible to attach textures and material to the surfaces of each object (Ledoux, Arroyo Ohori, et al., 2019). Thus, we define the aim of Category 4 to characterise attribute content. Considering that buildings are the most explicit theme and critical in several domains (Jaillot et al., 2020; Park & Guldmann, 2019; Willenborg et al., 2018), we only focus on building information in this work, e.g. the class, roof type, texture, measured height, etc. For future uses, criteria in Category 4 may be expanded depending on the specific context. For example, practitioners may determine extra criteria regarding the attributes of bridges, tunnels, tracks, city furniture, or other urban objects.

After designing criteria for the framework, a consistent scoring method is crucial to quantify the performance of each 3D city model, as well as indicate particular characteristics of datasets (Chi & Mak, 2021). The total score for this measurement is 47 points based on a 4-category scale, with 6 points for criteria around the data portal (Category 1), 15 points for basic information at Category 2, 11 points for thematic content (Category 3) and 15 points for attributes (Category 4). The score for each dataset could reflect its overall value, with higher scores generally indicating better performance. However, datasets with the same score may differ from each measurement category. For example, City A and City B score 25; City A receives 8 points in Category 2 and City B obtains 15 points in the same category. In order to present a standardised comparison, the scoring system intends to provide the result from two aspects. One is an overview of the measurement of included 3D city models by comparing their performance, and the other is investigating datasets' properties within each category.. As such, researchers and practitioners are able to select relevant parts in the dataset.

Our framework is not perfectly exhaustive, e.g. it may not include some attributes that are uncommon in practice, are not part of any major standard, and/or are developed for a very specific geographic or application context. However, if required by a particular use, it can be expanded to accommodate such information, e.g. building materials as per the CityGML Application Domain Extension for energy applications (Agugiaro et al., 2018), and verges as distinct features based on a national geoinformation standard (Brink et al., 2013). The framework is intended to be generic and flexible. It also enables practitioners to tailor it to specific use cases or geographical contexts. For example, in Category 4, we primarily regard building attributes as applications of 3D city models that primarily rely on buildings, e.g. the analysis of urban energy consumption (Biljecki et al., 2015). However, for further adoption, it also allows users to choose additional criteria in a specific context. For example, for research on the urban heat island effect, a set of attributes of other features can be added, e.g. trees, such as trunk height, vegetation canopy (Park et al., 2021).

Table 1: The 4-category framework with criteria.

Category	Criterion
(1) Data Portal	1C1 – Does the dataset have a dedicated website? 1C2 – Is there a web viewer in 3D? 1C3 – Is there near real-time information in the viewer? 1C4 – Is it available in local language? 1C5 – Is it available in English? 1C6 – Is there a way to leave feedback?
(2) Basic Information	2C1 – Is it a structured 3D city (information) model? 2C2 – Is it a 3D mesh model? 2C3 – Is it downloadable? 2C4 – Is it free of charge? 2C5 – Is it available to download without registration? 2C6 – Is it available to download in more than one format? 2C7 – Is it generated using open data standard? 2C8 – Is it openly licensed? 2C9 – Does it provide metadata? 2C10 – Is it recently published (within latest 5 years)? 2C11 – Has it been updated? 2C12 – Does it have a plan to update? 2C13 – Does it keep historical datasets? 2C14 – Does it include more than one level of detail (LoD)? 2C15 – Does it cover the entire jurisdiction?
(3) Thematic Content	3C1 – Are buildings modelled with semantically differentiated surfaces? 3C2 – Does it contain bridges? 3C3 – Does it contain land use? 3C4 – Does it contain terrain? 3C5 – Does it contain roads? 3C6 – Does it contain tunnels? 3C7 – Does it contain tracks? 3C8 – Does it contain city furniture? 3C9 – Does it contain vegetation? 3C10 – Does it contain individual trees? 3C11 – Does it contain water bodies?
(4) Attribute Content	4C1 – Does it contain the postal code? 4C2 – Are buildings with texture? 4C3 – Does it contain the ID of buildings? 4C4 – Does it contain the year of construction? 4C5 – Does it contain the address of buildings? 4C6 – Does it contain the function of buildings? 4C7 – Does it contain the height of buildings? 4C8 – Does it contain the volume of buildings? 4C9 – Does it contain the number of storeys? 4C10 – Does it contain the area of walls? 4C11 – Does it contain the area of roofs? 4C12 – Does it contain the type of roofs? 4C13 – Does it contain the area of grounds? 4C14 – Does it contain gross floor areas? 4C15 – Does it contain materials of buildings?

4 Implementation and Results

4.1 Inclusion Criteria

In this section, we put the framework into use by evaluating scores of datasets. While the framework is widely applicable, in our implementation, we define four criteria to filter datasets that we consider. First, datasets should be open. The Open Definition[5] defines open data as freely accessible and can be used and manipulated for any purpose. In addition, the openness of downloads is also related to availability, considering the operation of further analyses. Therefore, a 3D city model available on the Internet but requires payment or registration, or cannot be downloaded should not be included. In some cases, datasets are viewable but difficult to download. For example, some datasets require permission from the provider to receive a download link, which may take longer to process or be denied. Some require complex workflows to access them and are thus excluded being in contrast with openness. The second inclusion criterion is related to the thematic content: the dataset must contain buildings, as they are the main element of 3D city models and critical in several domains (Jaillot et al., 2020). That is, a 3D terrain model without buildings is technically a 3D city model, but may not be considered as such in practice (Kolbe et al., 2005). Coverage is the third consideration, referring to the administrative scope of the dataset. The dataset should cover the entire jurisdictional or regulatory boundary. While it is possible for a particular area to represent an entire region, it limits the ability to adapt further in practice (Goy et al., 2020). For example, each suburb in a city is in a different location, development and geomorphological condition. To use a particular suburb as a proxy for the city as a whole constitutes a coarse generalisation. Incomplete datasets ignore the urban environment while amplifying specific characteristics of particular areas. Therefore, the dataset must completely represent the jurisdictional area. Finally, the age of the 3D city model is the last inclusion criterion. Some users may be interested in historical data to analyse urban changes. However, age is another quality aspect, as many use cases require data with up-to-date information, whereas older data may be not applicable (Kaasenbrood, 2013; Zuiderwijk & Janssen, 2014). Therefore, we only consider datasets created within the last 10 years.

After applying the inclusion criteria, 40 city models are considered for evaluation in this work. To ensure standardisation and a valid comparison, they are all open government datasets that can be downloaded from government portals, including official national and municipal websites. Another reason for opting for open government datasets is the ease of obtaining them, without the need to purchase data or solicit access to non-open instances.

4.2 Description of one particular city: Linz (Austria)

To better understand how the framework works, we first single out a specific city to demonstrate the implementation in more detail. The authoritative 3D dataset of the

[5] https://opendefinition.org

City of Linz in Austria is considered an example to describe the properties of its 3D city model and the result of measurement (Figure 1). Overall, Linz has gained 22 points in total with 4 points, 11 points, 2 points and 5 points, respectively, in each category, which we describe in detail in the continuation.

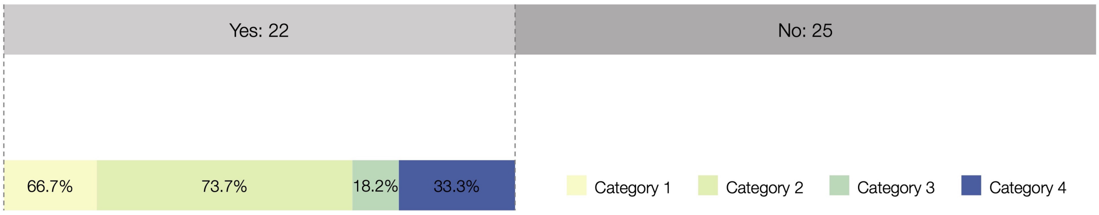

Figure 1: Visual representation of performance of Linz's 3D city model. The top plot indicates the share of criteria that are satisfied, while the bottom one suggests the achieved criteria by category.

In terms of data portal and its openness (Category 1), Linz[6] has a dedicated website to integrate basic information of its 3D city model and provide a web viewer for visualisation (Figure 2). Concerning language, it is developed in the local language, which will satisfy many local use cases. However, the absence of English is also a barrier to international users who do not understand German. In the viewer, the public can explore buildings with or without texture in terrain models. Additionally, it has a search bar to locate a specific address and obtain information about the buildings, for example, building use and the creation date. The data portal also provides contact information, allowing the public to leave feedback and welcome suggestions.

Considering the description of Linz's city model (Category 2), it performs well, achieving 11/15 points. First, it is downloadable with an open license and free of charge. The data provider also makes the metadata available to the public in case there are other demands for use. Besides, there is no need to make additional registers or requests to retrieve the datasets. However, it is not available to download in more than one format (CityGML). The age of the dataset is another significant criterion that indicates whether the city model reflects the current status of environment and delivers up-to-date information. In theory, the latest 3D city model has more potential to provide accurate and high-quality data. Linz's city model was created in 2020 based on LoD2 (i.e. providing simple building and roof design), considered a recent one in our framework. Moreover, in terms of time intervals, regular data updates are crucial as a high frequency can improve the overall quality of the dataset by verifying information and completing missing data. Linz's city model has been updated once but without a publicly specified plan to maintain it. Notably, this city model has kept the historical datasets in separate releases, which in some ways enables practitioners to conduct a historical analysis of urban development and reflect the change of Linz.

[6] https://3d.linz.at/

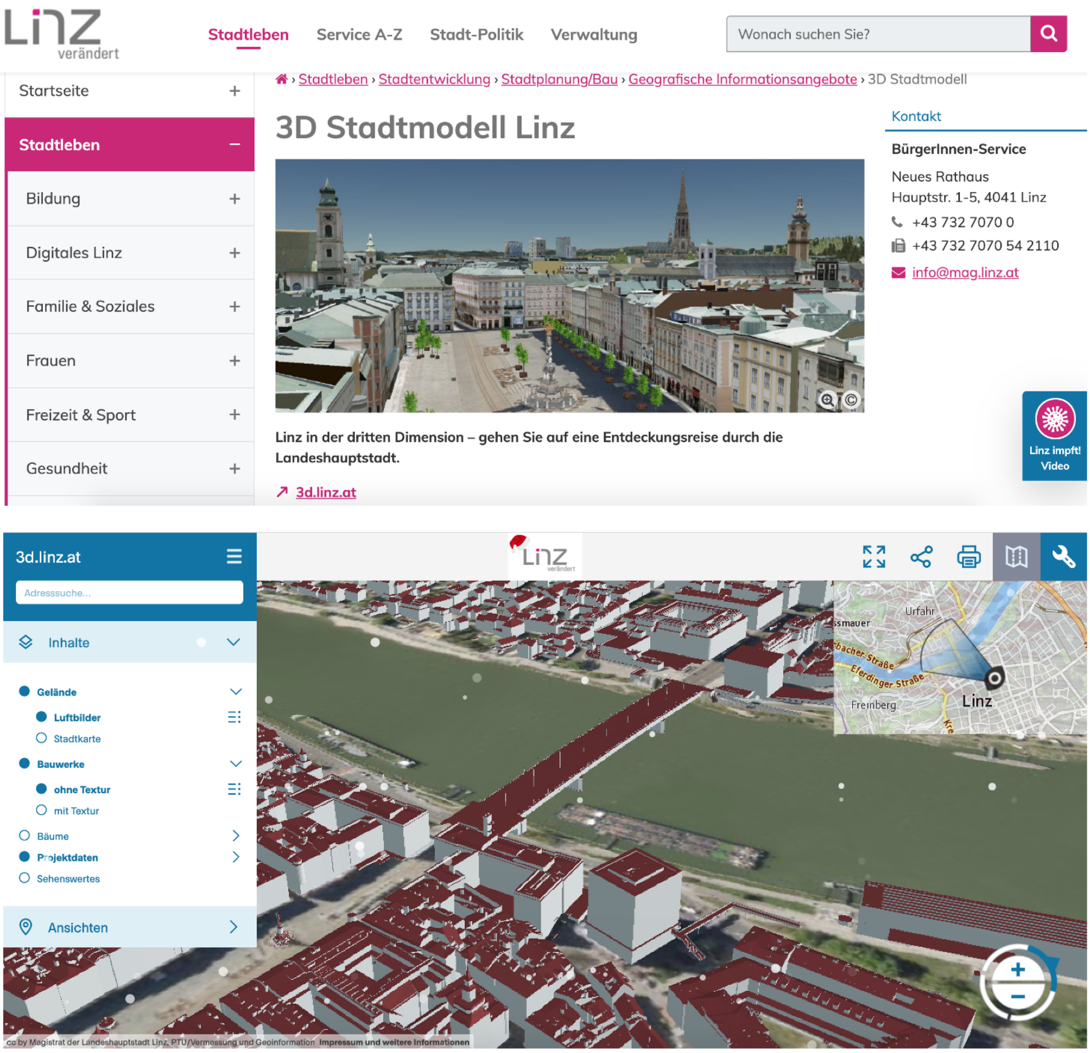

Figure 2: Screenshot of the government data portal for Linz's 3D city model, including the web viewer. The availability of such infrastructure is considered in our assessment framework. Image courtesy of Magistrat der Landeshauptstadt Linz, PTU/Vermessung und Geoinformation.

Linz's 3D city model gains 2 points among 11 questions when discussing its thematic content (Category 3). In addition to terrain and buildings, there are no other types of objects in the model, e.g. bridge, road, vegetation, or water. Regarding its attribute content (Category 4), Linz's city model scores 5/15 points. The dataset provides the ID of buildings, height, the area of walls and roofs, as well as the total area of each building (Figure 3). However, other detailed building information has not been included, e.g. the address of buildings, volume and the number of storeys.

The detailed examination and the resulting score indicate that the city model of Linz

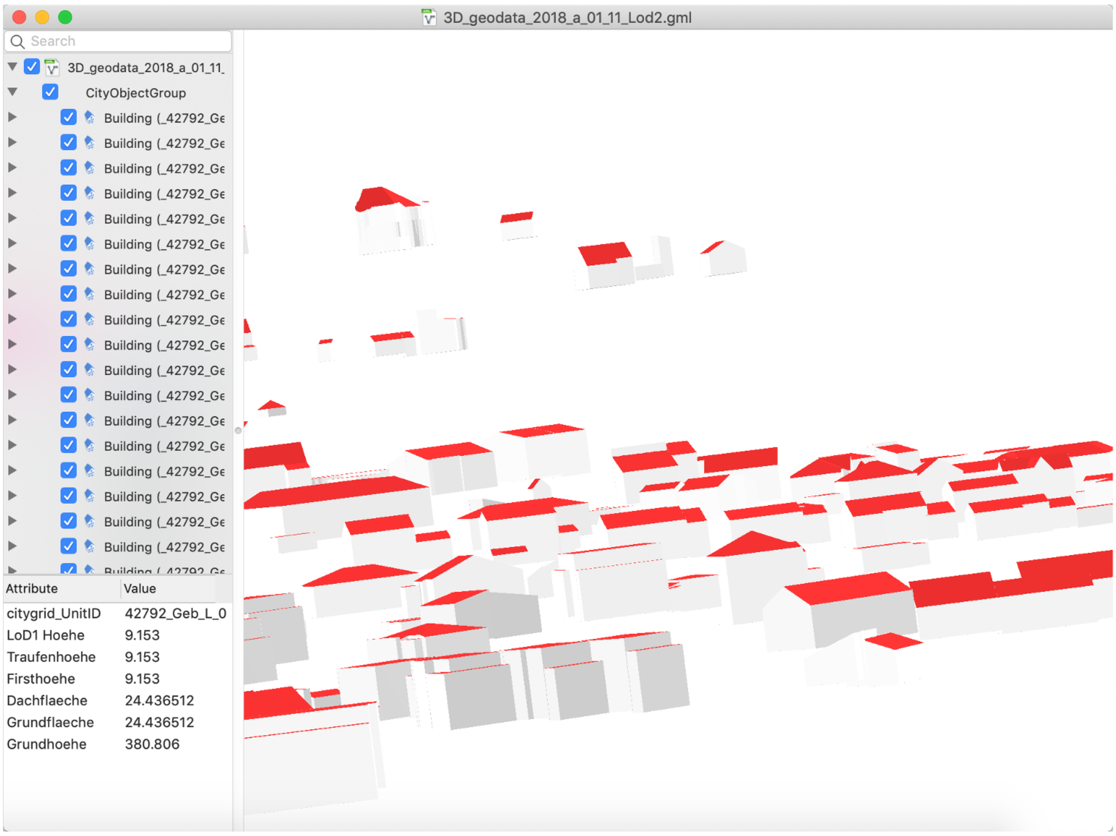

Figure 3: Visualisation of Linz's 3D dataset via Azul (Arroyo Ohori, 2020), including attribute content of each building. The semantic content of data plays an important role in the framework, while in the implementation, dozens of datasets have been inspected for the presence of particular attributes that are key to several use cases.

provides sufficient information to help stakeholders understand its characteristics and status. For example, this model is recently released and updated, which suggests that the information is new and up-to-date. Depending on specific purposes, some aspects will be more critical than others. For example, as Linz's city model has retained different versions of historical data, it is appropriate to track changes in the urban context. Considering the focus on content, Linz's city model is more suitable for building-related analysis than other domains.

4.3 Evaluating 3D city models in practice

The set of 40 datasets has been evaluated in the same fashion. A complete summary is illustrated in Figure 4. The result shows the scores (as an outcome) obtained in the framework with an indication of numerical values and a detailed breakdown of each category percentage of the individual dataset. The benchmarking of the included datasets suggests an overview of current development of 3D city models. Clearly, all datasets have significant differences in characteristics, reflecting their disconnected developments, lack of best practices, and heterogeneous purposes of local governments. It is also notable that some Japanese cities have similar performance overall. The reason is that the Ministry of Land, Infrastructure, Transport and Tourism (MLIT) in Japan has recently launched the PLATEAU project to create 3D city models of 66 cities in a unified manner, which explains the comparable performance of Japanese cities. As such, it indicates that some countries consider adopting 3D city models national and have citywide practices.

Another trend in evaluating datasets is the variation in scores between categories. Most city models perform well in Category 1, which implies a shared focus on model visualisation and accessibility. In Category 2, many datasets differ in the comprehensiveness of information provided the initiatives. For example, Zurich in Switzerland has the best performance receiving 14 out of 15 points (93.3%) under this category. During the in-depth analysis, we find that the local authority maintains a well-organised service to support the timeliness of information with regular update, as well as keep temporal versions of historical data. In contrast, Potsdam in Germany performs slightly deficiently receiving 7 points with a percentage of 46.7%. The latency of timeliness limits the potential to maintain its quality. The creation date is 2012, which is too old to reflect the latest urban environment. In addition, the update status has been static without a plan to update the dataset since published. With regard to thematic content (Category 3) and attributes (Category 4), many cities fail to include more object types and building information. Nevertheless, the 3D city model of Helsinki, as the best performer in implementing this framework, provides the most detailed building information. The richness of attribute content largely allows further analysis in specific scenarios. For example, one application of Helsinki city model is The Energy and Climate Atlas[7], supporting practitioners in investigating energy consumption and solar energy potential based on building-specific data.

To provide a meaningful context, we group datasets by age and update frequency to assess their characteristics in detail. For example, comparing an old dataset with a new one may not be relevant. With such stratification, we can also infer the progress of the development in the field, i.e. whether 3D city models that are created recently or regularly updated have made a leap or have particular characteristics over older counterparts. First, we split datasets into 2 groups by the criterion of 2C10 – Is it recently published (within the latest 5 years). Plotting total scores and scores as percentages in each measurement category clearly illustrates differences between newly published models and old ones (Figure 5). The result shows 6 datasets published over 5 years among all the 40 city models. The newer datasets receive more average points, indicating that they contain more comprehensive information. Notably, new models provide a detailed description of critical features (Category 2), more information about thematic content (Category 3), and building attributes (Category 4). It enables users to gain a better understanding of the 3D city model and then find out the most appropriate parts that they are interested in follow-up studies.

[7] https://kartta.hel.fi/3d/atlas/

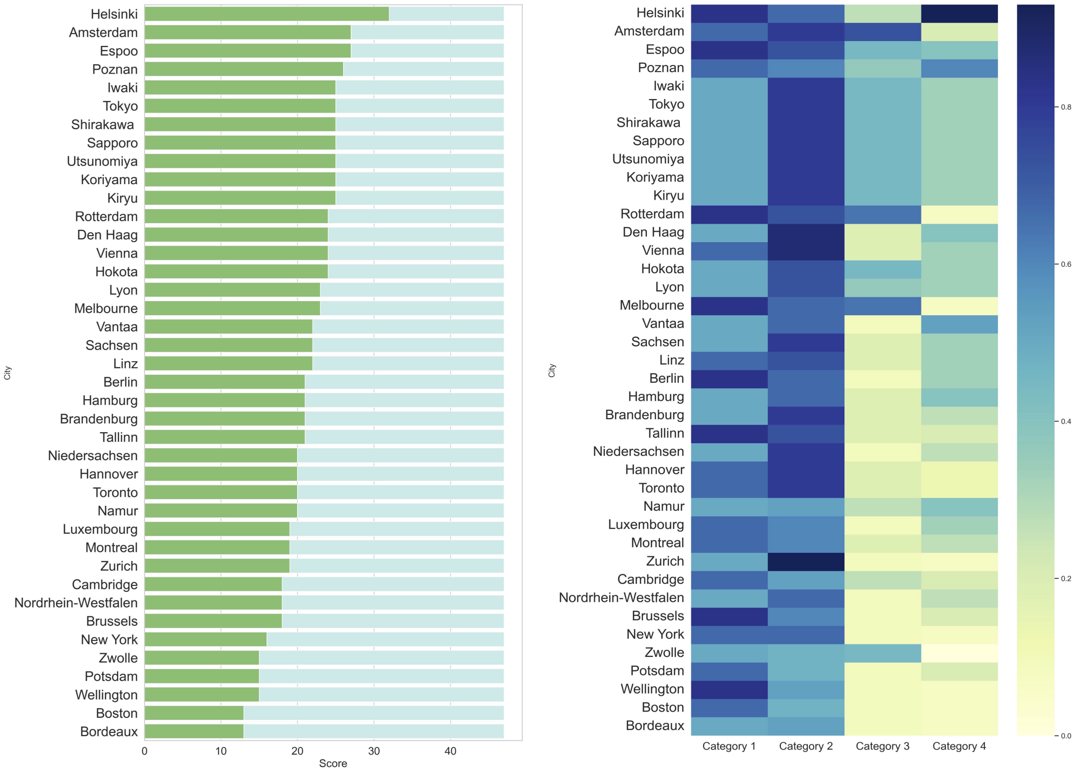

Figure 4: Results of implementation of the framework for the evaluated datasets. The left plot shows the scores obtained by each dataset, while the right one indicates the performance of each dataset within the four measurement categories (share of the fulfilled elements).

We next examine datasets by update frequency. Regarding update frequency, the framework offers two specific criteria to consider the status, including whether it has been updated at least once (2C11) and has the officially stated plans (2C12). All datasets are divided into three groups based on the update-related criteria: not updated (0 points for 2C11 and 2C12), irregularly updated (1 point for 2C11 or 2C12), and regularly updated (2 points for 2C11 and 2C12). A comprehensive assessment is shown in Figure 6. In general, the datasets without updates receive a lower score than the other two groups. Afterwards, their overall performance does not reflect a significant difference when diving into the irregularly updated and regularly updated group. The frequency of update has no notable impact on the content of 3D city models, which may imply that some updates only address a few minor modifications rather than enriching thematic content and building information. Although one may expect that datasets with a high frequency of update may potentially provide better information across many criteria, the measurement does not suggest such an advantage of regularly updated datasets. However, the timeliness of 3D city models plays a critical role in representing the physical environment and maintaining data quality and accuracy. By stratifying data by age and update frequency, we posit that the framework can provide the means to track the evolution of characteristics of 3D datasets over time.

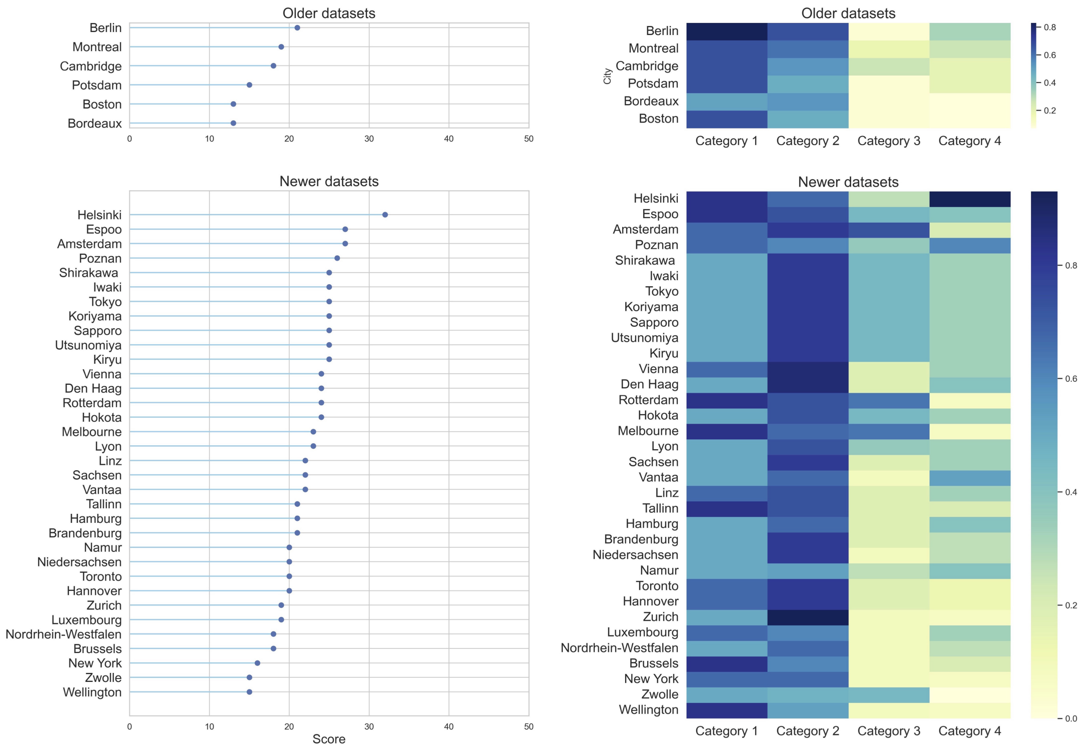

Figure 5: Performance among datasets of different age.

To better reflect the flexibility of the framework, we discuss some scenarios of the implementation in practice. Suppose a software developer is looking for a test 3D dataset for developing software, a task for which the location does not matter. In that case, they can use the framework to pick suitable datasets from around the world with particular characteristics they are interested in. When considering climate change, stakeholders can adopt this framework to identify which 3D city model is better for application in energy demand estimation studies, or what characteristics a city model should include. For example, solar potential analyses require detailed information such as roof type and availability of vegetation as it may cast shade affecting solar potential (Cheng et al., 2020; Saretta et al., 2020). In this case, users can make minor adaptions of the developed framework by adding more criteria on building properties and other relevant thematic objects, e.g. vegetation. For Category 1 and 2, the criteria can remain the same as they are about general information of datasets. The criteria under Category 3 and 4 could be slightly updated to suit this specific case. Users can also remove criteria that may be of little relevance in certain cases such as this one, e.g. bridges (3C2), roads (3C5), tunnels (3C6), tracks (3C7), and city furniture (3C8), and then add additional criteria of, for

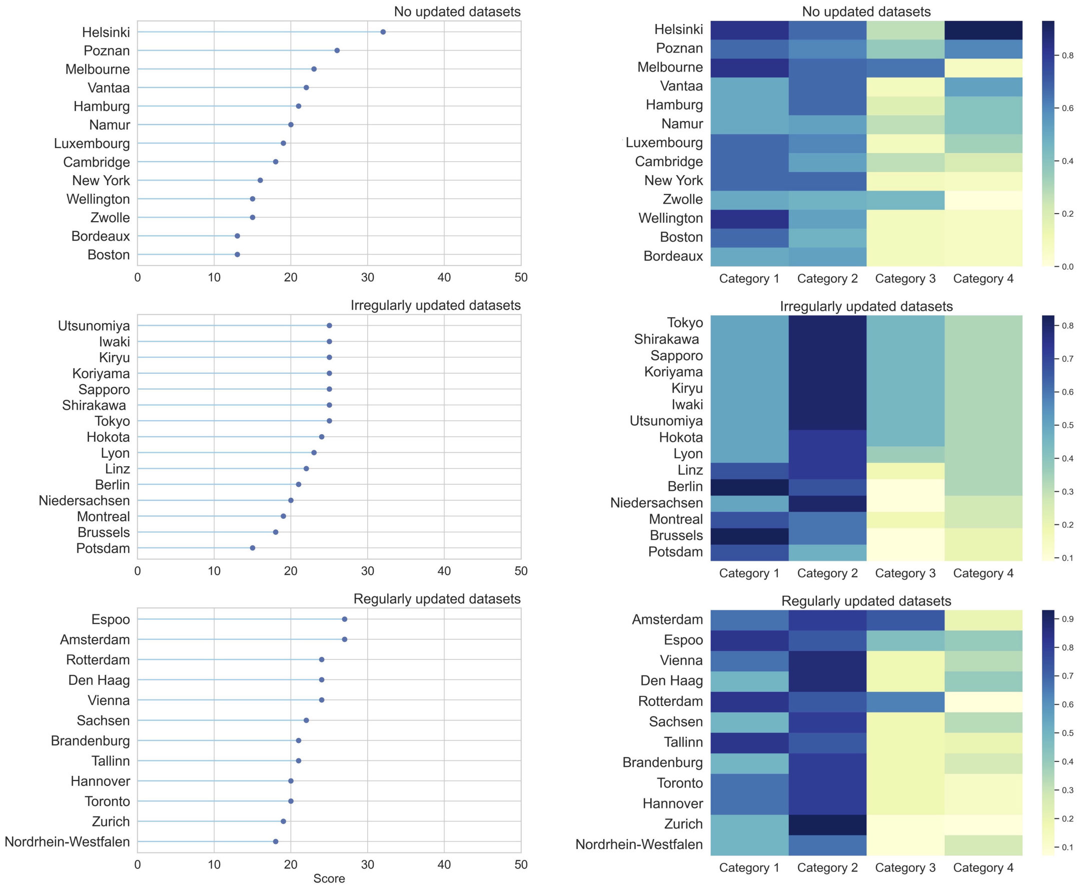

Figure 6: Performance among datasets of different update frequency.

example, wall height, wall aspect, vegetation canopy, and trunk height. For government users, this framework enables them to track best practices and drive the design of their future acquisitions. For other scenarios, users can also weigh each category to prioritise the objectives aligning with specific cases.

5 Discussion and Findings

The developed framework is a novel approach to assessing 3D city models, ranging from the portal level to detailed data characteristics. It balances a holistic and generic outlook and applicability to particular use cases. For example, it can be extended with further aspects by adding criteria to suit specific purposes, and conversely, criteria that may not be relevant can be omitted. As such, it will also contribute to customised uses depending on certain contexts. We believe that the framework, containing 47 criteria, manages to provide a holistic and robust instrument for a wide range of different stakeholders and use cases. At the same time, it provides the flexibility to further specify extra criteria based on particularities entailed by some use cases.

The results of the assessment lead to the conclusion that the differences among 3D city models are significant, while individual datasets show heterogeneous emphasis on different properties. Consequently, it confirms the capacity of the framework to capture the state of real-world 3D city models, as well as help researchers and practitioners understand the development of this field and common practices. Regarding properties, some 3D city models suffer from missing information in instances where different formats contain different degrees of richness. For instance, Espoo's 3D city model provides several formats for download. The Sketchup file delivers buildings in LoD1; however, the CityGML format contains more city objects in addition to buildings which is generated according to the open data standard, offering the potential for future interoperability. Moreover, it is acknowledged that various formats may satisfy different kinds of purposes (Biljecki & Arroyo Ohori, 2015; Herman & Reznik, 2015), but data providers still need to consider the consistency of information in each format and the adoption of open standard to ensure compatibility.

This work is the first endeavour to assess and compare 3D GIS datasets, together with associated aspects such as the means to obtain them (e.g. accessibility of data portal). Some limitations should be considered in future developments. One limitation is the selection of criteria. We consider the criteria based on relevant work in research on open government data and typical characteristics of datasets. However, there are a variety of criteria not selected in the framework. One example is data acquisition. The technique of collecting data may be relevant to some practitioners. Nevertheless, this aspect has not been included in the framework because it is often difficult to obtain and entails complexities (e.g. due to the increasing combination of acquisition techniques). For instance, a block 3D model derived by extruding building footprints to their height obtained from a point cloud may be more accurate than a detailed one that is derived from point clouds. One reason is that the former can be obtained from very accurate and high-quality data (e.g. super-dense point cloud). At the same time, the latter may be burdened by coarse and imperfect input datasets (e.g. coarse and inaccurate point cloud), entailing such criteria to be even misleading.

6 Conclusion

Following recent academic efforts in gauging open data and the characteristics of datasets in particular domains, we investigated the assessment of 3D city models and established a framework to evaluate and compare 3D GIS datasets, a first in the field. We believe that the work, named as '3D City Index', contributes to the community researching spatial data infrastructures. To the extent of our knowledge, there has been no related work in the domain of geographic information.

The developed framework contains four categories, which are: Data portal, Basic information, Thematic content, and Attribute content. Each category includes a set of criteria (in total we identify 47 criteria) to comprehensively emphasise different points of views (e.g. age and update frequency). Then, as a demonstration that doubles as an analysis of the worldwide state of open 3D geoinformation, a set of 40 authoritative open datasets has been assessed using the framework. Overall, our implementation contains about 2000 manually collected properties. Besides demonstrating the implementation of the framework, digging into the characteristics of 3D city models, there are valuable findings presented in this paper that may be useful by practitioners and researchers. The assessed 40 datasets present significant differences in each category. Such performances suggest heterogeneous approaches and different focuses of local governments. In addition, datasets that are newly released have more detailed properties than older datasets, especially the description of models. In contrast, the frequency of update has no significant impact on the richness of properties.

For future work, this framework can be applied to non-open datasets and those released by other organisations. It also allows for giving more weight to a particular aspect that may be more important than another in a specific case, allowing for a customisation for a certain context. Therefore, the value of this framework is not limited to this research, which a range of stakeholders can further realise to support different uses. In our future plan, we intend to create a website for the '3D City Index'. As such, it will be a sound way to maintain the sustainability of this framework and increase usability in practice. We believe the website will encourage the wide adoption of our framework by openly sharing it with the public and integrating discussion on its content. At the same time, we seek to continuously monitor the status of open 3D datasets and their characteristics. Such continuity will enable a better understanding of the trends and provide insight into developing 3D city models worldwide. Whether people have knowledge of 3D city models or are new to this field, they can find helpful information and adopt some parts in practical uses from the framework and the website.

Code and Data availability statement

The dataset and analysis that supports the findings of this study are available at https://doi.org/10.6084/m9.figshare.19221852.v8.

Competing interests

The authors report there are no competing interests to declare.

Acknowledgements

We gratefully acknowledge the data used in this research. We thank the members of the NUS Urban Analytics Lab for the discussions. This research is part of the projects (i) Large-scale 3D Geospatial Data for Urban Analytics, which is supported by the National University of Singapore under the Start Up Grant R-295-000-171-133; and (ii) Multi-scale Digital Twins for the Urban Environment: From Heartbeats to Cities, which is supported by the Singapore Ministry of Education Academic Research Fund Tier 1.

Notes on contributors

Binyu Lei is a PhD researcher in the Urban Analytics Lab at the National University of Singapore. She holds a Master degree in Urban Planning from the University of Melbourne.

Rudi Stouffs is Dean's Chair Associate Professor and Assistant Dean (Research) at the National University of Singapore. He leads the Architectural and Urban Prototyping lab, and is principal investigator in the Future Resilient Systems II and the FCL Global research programmes coordinated by the Singapore ETH Centre. He holds a PhD in Architecture from Carnegie Mellon University.

Filip Biljecki is an Assistant Professor at the National University of Singapore and the principal investigator of the NUS Urban Analytics Lab. He holds an MSc in Geomatics and a PhD in 3D GIS from the Delft University of Technology in the Netherlands.

References

Agugiaro, G., Benner, J., Cipriano, P., & Nouvel, R. (2018). The Energy Application Domain Extension for CityGML: enhancing interoperability for urban energy simulations. *Open Geospatial Data, Software and Standards*, *3*(1), 139. https://doi.org/10.1186/s40965-018-0042-y

Alam, N., Wagner, D., Wewetzer, M., Falkenhausen, J. v., Coors, V., & Pries, M. (2014). Towards Automatic Validation and Healing of Citygml Models for Geometric and Semantic Consistency. *Innovations in 3D geo-information sciences* (pp. 77–91). Springer. https://doi.org/10.5194/isprsannals-II-2-W1-1-2013

Arroyo Ohori, K. (2020). azul: A fast and efficient 3D city model viewer for macOS. *Transactions in GIS*, *24*(5), 1165–1184. https://doi.org/10.1111/tgis.12673

Attard, J., Orlandi, F., Scerri, S., & Auer, S. (2015). A systematic review of open government data initiatives. *Government information quarterly*, *32*(4), 399–418. https://doi.org/10.1016/j.giq.2015.07.006

Berners-Lee, T. (2006). Linked data. *Int. J. on Semantic Web and Information Systems*, *4*(2).

Biljecki, F., & Tauscher, H. (2019). Quality of BIM–GIS conversion. *ISPRS Ann. Photogramm. Remote Sens. Spatial Inf. Sci.*, *IV-4/W8*, 35–42. https://doi.org/10.5194/isprs-annals-iv-4-w8-35-2019

Biljecki, F., & Arroyo Ohori, K. (2015). Automatic semantic-preserving conversion between OBJ and CityGML. *UDMV15: Eurographics Workshop on Urban Data Modelling and Visualisation, Delft, The Netherlands, 23 November 2015; Authors version.* https://doi.org/10.2312/udmv.20151345

Biljecki, F., Ledoux, H., Du, X., Stoter, J., Soon, K. H., & Khoo, V. (2016). The most common geometric and semantic errors in CityGML datasets. *ISPRS Annals* *of Photogrammetry, Remote Sensing & Spatial Information Sciences*, *4*. https://doi.org/10.5194/isprs-annals-IV-2-W1-13-2016

Biljecki, F., Lim, J., Crawford, J., Moraru, D., Tauscher, H., Konde, A., Adouane, K., Lawrence, S., Janssen, P., & Stouffs, R. (2021). Extending CityGML for IFC-sourced 3D city models. *Automation in Construction*, *121*, 103440. https://doi.org/10.1016/j.autcon.2020.103440

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*(4), 2842–2889. https://doi.org/10.3390/ijgi4042842

Bognár, Á., Loonen, R. C., & Hensen, J. L. (2021). Calculating solar irradiance without shading geometry: A point cloud-based method. *Journal of Building Performance Simulation*, *14*(5), 480–502. https://doi.org/10.1080/19401493.2021.1971765

Borzacchiello, M. T., & Craglia, M. (2012). The Impact on Innovation of Open Access to Spatial Environmental Information: A Research Strategy. *International Journal of Technology Management*, *60*(1-2), 114–129. https://doi.org/10.1504/IJTM.2012.049109

Braunschweig, K., Eberius, J., Thiele, M., & Lehner, W. (2012). The State of Open Data Limits of Current Open Data Platforms. *Conference Proceedings Eorld Wide Web Conference*.

Brink, L. v. d., Stoter, J., & Zlatanova, S. (2013). UML-Based Approach to Developing a CityGML Application Domain Extension. *Transactions in GIS*, *17*(6), 920–942. https://doi.org/10.1111/tgis.12026

Cheng, L., Zhang, F., Li, S., Mao, J., Xu, H., Ju, W., Liu, X., Wu, J., Min, K., Zhang, X., & Li, M. (2020). Solar energy potential of urban buildings in 10 cities of China. *Energy*, *196*, 117038. https://doi.org/https://doi.org/10.1016/j.energy.2020.117038

Chi, Y. L., & Mak, H. W. L. (2021). From Comparative and Statistical Assessments of Liveability and Health Conditions of Districts in Hong Kong towards Future City Development. *Sustainability*, *13*(16), 8781. https://doi.org/10.3390/su13168781

Conradie, P., & Choenni, S. (2012). Exploring process barriers to release public sector information in local government. *Proceedings of the 6th international conference on theory and practice of electronic governance*, 5–13. https://doi.org/10.1145/2463728.2463731

Coors, V., Betz, M., & Duminil, E. (2020). A Concept of Quality Management of 3D City Models Supporting Application-Specific Requirements. *PFG – Journal of Photogrammetry, Remote Sensing and Geoinformation Science*, 1–12. https://doi.org/10.1007/s41064-020-00094-0

Degbelo, A. (2021). FAIR geovisualizations: definitions, challenges, and the road ahead. *International Journal of Geographical Information Science*, 1–41. https://doi.org/10.1080/13658816.2021.1983579

Dukai, B., Peters, R., Vitalis, S., van Liempt, J., & Stoter, J. (2021). Quality assessment of a nationwide data set containing automatically reconstructed 3D building models. *The International Archives of the Photogrammetry, Remote Sensing and*

Spatial Information Sciences, *XLVI-4/W4-2021*, 17–24. https://doi.org/10.5194/isprs-archives-xlvi-4-w4-2021-17-2021

Dukai, B., Peters, R., Wu, T., Commandeur, T., Ledoux, H., Baving, T., Post, M., van Altena, V., van Hinsbergh, W., & Stoter, J. (2020). Generating, storing, updating and disseminating a countrywide 3D model. *The International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences*, *44*, 27–32. https://doi.org/10.5194/isprs-archives-XLIV-4-W1-2020-27-2020

Ennafii, O., Bris, A. L., Lafarge, F., & Mallet, C. (2019). A Learning Approach to Evaluate the Quality of 3D City Models. *Photogrammetric Engineering & Remote Sensing*, *85*(12), 865–878. https://doi.org/10.14358/pers.85.12.865

Farhang, L., Bhatia, R., Scully, C. C., Corburn, J., Gaydos, M., & Malekafzali, S. (2008). Creating tools for healthy development: case study of San Francisco's Eastern Neighborhoods Community Health Impact Assessment. *Journal of public health management and practice*, *14*(3), 255–265. https://doi.org/10.1097/01.PHH.0000316484.72759.7b

Gil, J. (2020). City Information Modelling: A Conceptual Framework for Research and Practice in Digital Urban Planning. *Built Environment*, *46*(4), 501–527. https://doi.org/10.2148/benv.46.4.501

Goy, S., Maréchal, F., & Finn, D. (2020). Data for Urban Scale Building Energy Modelling: Assessing Impacts and Overcoming Availability Challenges. *Energies*, *13*(16), 4244. https://doi.org/10.3390/en13164244

Gröger, G., Kolbe, T. H., Nagel, C., & Häfele, K.-H. (2012). OGC city geography markup language (CityGML) encoding standard.

Herman, L., & Reznik, T. (2015). 3D web visualization of environmental information-integration of heterogeneous data sources when providing navigation and interaction. *The International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences*, *40*(3), 479. https://doi.org/10.5194/isprsarchives-XL-3-W3-479-2015

Hossain, M. A., Dwivedi, Y. K., & Rana, N. P. (2016). State-of-the-art in open data research: Insights from existing literature and a research agenda. *Journal of organizational computing and electronic commerce*, *26*(1-2), 14–40. https://doi.org/10.1080/10919392.2015.1124007

Jaillot, V., Servigne, S., & Gesquière, G. (2020). Delivering time-evolving 3D city models for web visualization. *International Journal of Geographical Information Science*, *34*(10), 2030–2052. https://doi.org/10.1080/13658816.2020.1749637

Julin, A., Jaalama, K., Virtanen, J.-P., Pouke, M., Ylipulli, J., Vaaja, M., Hyyppä, J., & Hyyppä, H. (2018). Characterizing 3D city modeling projects: towards a harmonized interoperable system. *ISPRS International Journal of Geo-Information*, *7*(2), 55.

Kaasenbrood, M. (2013). An exploration of the use of open government data by private organisations: Contributing to the improvement of governmental policies by examining the current use of open government data by private organisations in the Netherlands.

Kedron, P., Li, W., Fotheringham, S., & Goodchild, M. (2021). Reproducibility and replicability: Opportunities and challenges for geospatial research. *International Journal of Geographical Information Science*, *35*(3), 427–445. https://doi.org/10.1080/13658816.2020.1802032

Kolbe, T. H. (2009). Representing and Exchanging 3D City Models with CityGML. *3d geo-information sciences* (pp. 15–31). Springer. https://doi.org/10.1007/978-3-540-87395-2_2

Kolbe, T. H., & Donaubauer, A. (2021). Semantic 3D City Modeling and BIM. *Urban informatics* (pp. 609–636). Springer. https://doi.org/10.1007/978-981-15-8983-6_34

Kolbe, T. H., Gröger, G., & Plümer, L. (2005). CityGML: Interoperable access to 3D city models. *Geo-information for disaster management* (pp. 883–899). Springer.

Kundra, V. (2011). Federal cloud computing strategy.

Labetski, A., Kumar, K., Ledoux, H., & Stoter, J. (2018). A metadata ADE for CityGML. *Open Geospatial Data, Software and Standards*, *3*(1). https://doi.org/10.1186/s40965-018-0057-4

Labetski, A., Vitalis, S., Biljecki, F., Ohori, K. A., & Stoter, J. (2022). 3D building metrics for urban morphology. *International Journal of Geographical Information Science*, *0*(0), 1–32. https://doi.org/10.1080/13658816.2022.2103818

Ledoux, H. (2018). val3dity: validation of 3D GIS primitives according to the international standards. *Open Geospatial Data, Software and Standards*, *3*, 1. https://doi.org/10.1186/s40965-018-0043-x

Ledoux, H., Arroyo Ohori, K., Kumar, K., Dukai, B., Labetski, A., & Vitalis, S. (2019). CityJSON: A compact and easy-to-use encoding of the CityGML data model. *Open Geospatial Data, Software and Standards*, *4*(1), 1–12.

Ledoux, H., Biljecki, F., Dukai, B., Kumar, K., Peters, R., Stoter, J., & Commandeur, T. (2021). 3dfier: automatic reconstruction of 3D city models. *Journal of Open Source Software*, *6*(57), 2866. https://doi.org/10.21105/joss.02866

Ledoux, H., Ohori, K. A., Kumar, K., Dukai, B., Labetski, A., & Vitalis, S. (2019). CityJSON: A compact and easy-to-use encoding of the CityGML data model. *Open Geospatial Data, Software and Standards*, *4*(1). https://doi.org/10.1186/s40965-019-0064-0

Liang, J., Gong, J., Xie, X., & Sun, J. (2020). Solar3D: An Open-Source Tool for Estimating Solar Radiation in Urban Environments. *ISPRS International Journal of Geo-Information*, *9*(9), 524. https://doi.org/10.3390/ijgi9090524

Lim, J., Janssen, P., & Biljecki, F. (2020). Visualising detailed CityGML and ADE at the building scale. *Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci.*, *XLIV-4/W1-2020*, 83–90. https://doi.org/10.5194/isprs-archives-xliv-4-w1-2020-83-2020

Mak, H. W. L., & Lam, Y. F. (2021). Comparative assessments and insights of data openness of 50 smart cities in air quality aspects. *Sustainable Cities and Society*, *69*, 102868. https://doi.org/10.1016/j.scs.2021.102868

Noardo, F., Wu, T., Ohori, K. A., Krijnen, T., & Stoter, J. (2022). IFC models for semi-automating common planning checks for building permits. *Automation in Construction*, *134*, 104097. https://doi.org/10.1016/j.autcon.2021.104097

Palliwal, A., Song, S., Tan, H. T. W., & Biljecki, F. (2021). 3D city models for urban farming site identification in buildings. *Computers, Environment and Urban Systems*, *86*, 101584. https://doi.org/10.1016/j.compenvurbsys.2020.101584

Park, Y., & Guldmann, J.-M. (2019). Creating 3D city models with building footprints and LIDAR point cloud classification: A machine learning approach. *Computers, environment and urban systems*, *75*, 76–89. https://doi.org/https://doi.org/10.1016/j.compenvurbsys.2019.01.004

Park, Y., Guldmann, J.-M., & Liu, D. (2021). Impacts of tree and building shades on the urban heat island: Combining remote sensing, 3D digital city and spatial regression approaches. *Computers, Environment and Urban Systems*, *88*, 101655. https://doi.org/10.1016/j.compenvurbsys.2021.101655

Peters, R., Dukai, B., Vitalis, S., van Liempt, J., & Stoter, J. (2022). Automated 3D Reconstruction of LoD2 and LoD1 Models for All 10 Million Buildings of the Netherlands. *Photogrammetric Engineering & Remote Sensing*, *88*(3), 165–170. https://doi.org/10.14358/pers.21-00032r2

Sachs, L. (2012). *Applied statistics: A handbook of techniques*. Springer Science & Business Media.

Saretta, E., Bonomo, P., & Frontini, F. (2020). A calculation method for the BIPV potential of Swiss façades at LOD2.5 in urban areas: A case from Ticino region. *Solar Energy*, *195*, 150–165. https://doi.org/https://doi.org/10.1016/j.solener.2019.11.062

Schröter, K., Lüdtke, S., Redweik, R., Meier, J., Bochow, M., Ross, L., Nagel, C., & Kreibich, H. (2018). Flood loss estimation using 3D city models and remote sensing data. *Environmental Modelling & Software*, *105*, 118–131. https://doi.org/10.1016/j.envsoft.2018.03.032

Shahat, E., Hyun, C. T., & Yeom, C. (2021). City Digital Twin Potentials: A Review and Research Agenda. *Sustainability*, *13*(6). https://doi.org/10.3390/su13063386

Shannon, J., & Walker, K. (2018). Opening GIScience: A process-based approach. *International Journal of Geographical Information Science*, *32*(10), 1911–1926. https://doi.org/10.1080/13658816.2018.1464167

Stoter, J., Peters, R., Commandeur, T., Dukai, B., Kumar, K., & Ledoux, H. (2020). Automated reconstruction of 3D input data for noise simulation. *Computers, Environment and Urban Systems*, *80*, 101424. https://doi.org/10.1016/j.compenvurbsys.2019.101424

Strathern, M. (2000). The Tyranny of Transparency. *British educational research journal*, *26*(3), 309–321.

Susha, I., Zuiderwijk, A., Janssen, M., & Grönlund, Å. (2015). Benchmarks for Evaluating the Progress of Open Data Adoption: Usage, Limitations, and Lessons Learned. *Social science computer review*, *33*(5), 613–630. https://doi.org/10.1177/0894439314560852

Tang, L., Ying, S., Li, L., Biljecki, F., Zhu, H., Zhu, Y., Yang, F., & Su, F. (2020). An application-driven LOD modeling paradigm for 3D building models. *ISPRS Journal of Photogrammetry and Remote Sensing*, *161*, 194–207. https://doi.org/10.1016/j.isprsjprs.2020.01.019

Ubaldi, B. (2013). Open government data: Towards empirical analysis of open government data initiatives.

Vetrò, A., Canova, L., Torchiano, M., Minotas, C. O., Iemma, R., & Morando, F. (2016). Open data quality measurement framework: Definition and application to Open Government Data. *Government Information Quarterly*, *33*(2), 325–337. https://doi.org/10.1016/j.giq.2016.02.001

Virtanen, J.-P., Jaalama, K., Puustinen, T., Julin, A., Hyyppä, J., & Hyyppä, H. (2021). Near Real-Time Semantic View Analysis of 3D City Models in Web Browser. *ISPRS International Journal of Geo-Information*, *10*(3), 138. https://doi.org/10.3390/ijgi10030138

Vitalis, S., Arroyo Ohori, K., & Stoter, J. (2020). CityJSON in QGIS: Development of an open-source plugin. *Transactions in GIS*, *24*(5), 1147–1164. https://doi.org/10.1111/tgis.12657

Wagner, D., Alam, N., Wewetzer, M., Pries, M., & Coors, V. (2015). Methods for geometric data validation of 3D city models. *Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci.*, *XL-1-W5*, 729–735. https://doi.org/10.5194/isprsarchives-xl-1-w5-729-2015

Willenborg, B., Sindram, M., & Kolbe, T. H. (2018). Applications of 3D city models for a better understanding of the built environment. *Trends in Spatial Analysis and Modelling*, 167–191. https://doi.org/10.1007/978-3-319-52522-8_9

Wilson, J. P., Butler, K., Gao, S., Hu, Y., Li, W., & Wright, D. J. (2021). A Five-Star Guide for Achieving Replicability and Reproducibility When Working with GIS Software and Algorithms. *Annals of the American Association of Geographers*, *111*(5), 1–7. https://doi.org/10.1080/24694452.2020.1806026

Wu, A. N., & Biljecki, F. (2022). GANmapper: geographical data translation. *International Journal of Geographical Information Science*, *36*, 1394–1422. https://doi.org/10.1080/13658816.2022.2041643

Wysocki, O., Schwab, B., Hoegner, L., Kolbe, T. H., & Stilla, U. (2021). Plastic Surgery for 3D City Models: a Pipeline for Automatic Geometry Refinement and Semantic Enrichment. *ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, *V-4-2021*, 17–24. https://doi.org/10.5194/isprs-annals-v-4-2021-17-2021

Zuiderwijk, A., & Janssen, M. (2014). Barriers and Development Directions for the Publication and Usage of Open Data: A Socio-Technical View. *Open government* (pp. 115–135). Springer. https://doi.org/10.1007/978-1-4614-9563-5_8