QUALITY OF BIM–GIS CONVERSION

Filip Biljecki [1,*], Helga Tauscher [2]

[1] National University of Singapore, Singapore – filip@nus.edu.sg
[2] Dresden University of Applied Sciences, Dresden, Germany – helga.tauscher@htw-dresden.de

KEY WORDS: IFC, CityGML, quality, errors, topology, BIM, GIS, validation

ABSTRACT:

Much work has been done on quality of geoinformation and interoperability between BIM and GIS. However, the intersection of the two — quality control of the conversion between BIM and GIS — remains uncharted. This discussion paper, based on empirical results, is one of the first steps towards mapping out a framework on errors and quality control in the context of BIM–GIS interoperability. In our work we focus on the conversion from IFC to CityGML, identifying several systematic errors potentially common and/or exclusive to the context of BIM–GIS conversion. Besides exposing several faults pertaining to IFC-sourced 3D city models, we discuss their taxonomy and their potential impact when engaged in applications. This paper is also relevant with respect to the growing popularity of conversion between IFC and CityGML, potentially aiding others to avoid many of the errors that can occur in the process and establishing directions to set up a benchmark to assess the performance of the interoperability workflows.

1. INTRODUCTION

The two topics in 3D GIS that have recently been a subject of a substantial number of research projects and publications are (i) data quality and (ii) integration with BIM. For example, research has been done on defining a set of rules for the geometry of 3D city models and developing software validators ensuring those. On the other hand, IFC models are becoming increasingly popular in the geospatial domain as sources of highly detailed datasets and many research groups around the world have been working to leverage them by developing integration and conversion procedures between data models. A comprehensive literature review on both topics is presented in Section 2.

Notably missing in literature is the intersection of the two topics — the quality aspect of the conversion between BIM and GIS data. We partially bridge that void with this paper by converging the two through conveying our experience from a recently completed project on IFC-to-CityGML conversion. We focus on the quality aspect of the conversion results, with the aim to foster discussions leading to a framework on the quality assessment of the BIM and GIS interoperability.

In the paper we first list common errors we have encountered in the output CityGML data during the development of our conversion methodology, and also include experiences of other researchers who have documented these. While many errors and inconsistencies that our and other papers expose are not new to the literature (e.g. overlapping geometries), there is a subset of errors that we deem specific or at least common in the BIM–GIS environment, necessitating discussions. Thus we highlight these errors, and work on classifying them and discussing the potential impact they have when engaged in applications.

The impact in applications is important for BIM–GIS interoperability. Most papers on quality of geoinformation are either generic or focus on errors relevant to a single application (e.g. a particular spatial analysis). However, because BIM–GIS interoperability is often constituted to support a specific application, e.g. microscale flood damage assessment (Amirebrahimi et al., 2015), we deem that it is important to define errors based on specific application requirements and to consider the severity of errors by examining their impact when used in an application.

The quality issues we have encountered during our project represent quality aspects that might be important to pay attention on in similar projects. Although this paper is shaped after our experiences in IFC-to-CityGML conversion, much of the paper may also be applicable in the reverse direction, conversion from and to other formats, and model synchronization; and thus – it should be viewed in broader context of the topic of the BIM–GIS interoperability. Furthermore, we hope that our paper will help developing a mechanism for validation and benchmarking of a BIM to GIS conversion.

2. BACKGROUND

2.1 BIM and GIS interoperability

There are not many other recent topics in 3D GIS that have been investigated so intensively as the integration with BIM (Liu et al., 2017; Zhao et al., 2019; Wang et al., 2019; Kang, Hong). Existing work largerly focusses on bridging the buildingSmart standard Industry Foundation Classes (IFC) and the Open Geospatial Consortium (OGC) standard CityGML (Isikdag, Zlatanova; Jusuf et al., 2017; Floros et al., 2018; Kang, Hong; Stouffs et al., 2018; Donkers et al., 2016), the two prominent data models in the architecture, engineering, and construction (buildingSMART, 2016; ISO, 2013) domain and the geospatial worlds (Open Geospatial Consortium, 2012; Gröger, Plümer), respectively. However, there still remain many challenges in bridging the two worlds (Kumar et al., 2019).

It is worth observing that most work has been developed for a specific application, e.g. fire-fighting simulations (Chen et al., 2014), subway tunnel design (Borrmann et al., 2015), noise mapping (Deng et al., 2016a), and view analyses (Rafiee et al., 2014). This is a relevant observation for this work, because we shape our discussion and classification of errors by their impact on applications. This is not a new point of view in 3D GIS (Sargent et al., 2007), as data quality in GIS is often assessed by its

* Corresponding author

fitness for use for a specific application (Devillers et al., 2007), and in fact the existence of an error can only be defined against an expectation or requirement.

In the context of this paper, it is also important to note some characteristics of CityGML datasets sourced from IFC. First, the scope of BIM leads to datasets with a very high level of detail compared to typical GIS data (Stouffs et al., 2018). Second, researchers developing workflows for the conversion often modify the target schema, e.g. extend the CityGML standard data model (Biljecki et al., 2018).

Despite the abundance of papers on this topic, quality is rarely discussed. In fact, searching for terms such as 'error' in most papers does not yield a single occurrence. On the other hand, it should be noted that some researchers not only point out errors, but also manage to repair them on-the-fly during the conversion (Donkers et al., 2016).

2.2 Quality of 3D city models

Quality assessment of 3D city models has become a prominent subject in the recent a few years, resulting in developing guidelines and software for validation (Ledoux, 2013; Wagner et al., 2015; OGC, 2016). Most of the work is generic (application-agnostic), with some exceptions (Ellul et al., 2016).

A large-scale assessment of CityGML datasets available in practice has been carried out by Biljecki et al. (2016). They list most common errors found across 37 publicly available datasets (comprising 3.6 million buildings). We have encountered some of these errors in our work, but they are not all specific to BIM—GIS interoperability. However, three IFC-specific errors have been highlighted by the researchers: modelling volumetric surfaces in CityGML 2.0, omission of features such as walls, and misclassification of semantic surfaces (e.g. `GroundSurface` instead of `FloorSurface`). Another observation from this large-scale assessment is that datasets modelled at a higher level of detail tend to have more errors. Because BIM nominally provides a source of highly detailed datasets, IFC-sourced 3D city models would naturally be prone to more errors.

2.3 Our project IFC2CityGML

To give more context to this work, we will briefly describe the efforts from which we obtain the experience laid out in this paper. A particularity of our IFC–CityGML conversion method is that it is configurable through rules (e.g. with respect to the spatio-semantic paradigms used) and thus during our developments we have produced a variety of CityGML output (Figures 1 and 2): constructive elements, boundary surfaces, and floor plans (the new LoD0 concept (Löwner et al., 2016)). Subsequently, the different flavours of output datasets expose a variety of errors.

Preliminary papers about the project are available (Stouffs et al., 2018; Konde et al., 2018; Tauscher, Stouffs; Lim et al., 2019). Also, an ADE has been developed to preserve rich information from IFC. This is an important aspect of the work because the quality examination should cover to the extended data model as well.

2.4 Relevant efforts

Although to the extent of our knowledge there is no work specifically focusing on the quality aspect of BIM–GIS interoperability, some of the papers describing efforts in developing conversion mechanisms between different data models (Section 2.1) briefly mention issues and errors that were exposed during the development (e.g. see (Lilis et al., 2016)). We list some of these errors later in the paper. There is a general agreement in literature that the output dataset can only be as good as the input dataset, and that the conversion may induce errors, but the papers mostly fall short in elaborating on details.

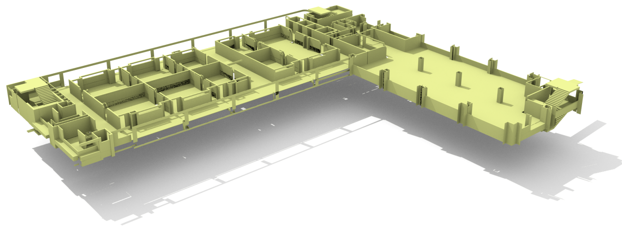

Figure 1. CityGML dataset obtained from IFC (storey of one building) with the approach developed within our project. Source of the input IFC dataset: BCA Singapore.

There is some related work beneficial to mention. Arroyo Ohori et al. (2017) point out geometric and topological issues in the integration of GIS and BIM data. For example, the researchers reveal that self-intersections and intersections are quite common. Furthermore, Arroyo Ohori et al. (2018) and Lilis et al. (2015) highlight that many IFC datasets in practice have topological errors, which are essential to repair before proceeding them into conversion workflows (for examples of diverse errors found in IFC and efforts to identify them see (Lee et al., 2015)).

Among other tasks, the ongoing ISPRS/EuroSDR GeoBIM benchmark 2019 (Noardo et al., 2019), investigates software tools for conversion of data from IFC to CityGML and vice versa. The benchmark investigates different aspects of the conversion such as performance. However, it does not focus much on quality of the data and since the project is ongoing the results are not available yet.

Sun et al. (2018) discuss preliminary plans on evaluating quality (e.g. positional accuracy) in their project integration of BIM and geospatial data. Similarly the project is ongoing, thus only limited information is available.

3. METHODOLOGY AND TOOLS

The methodology of our work is to list errors that we have encountered during the conversion from IFC to CityGML (Section 4), and to discuss the establishment of a framework for quality assessment of BIM–GIS interoperability taking into account the impact quality issues may have in applications (Section 5). Identifying the errors in the CityGML output also enabled us to go back and forth to the input IFC dataset and the conversion workflow exposing faults that caused them and helped us to fix the issues.

We have used multiple approaches and tools to detect the errors. Some errors became obvious during manual inspection of the source and visual rendering. However, although an output dataset may visually look correct and consistent, it may be erroneous (Lilis et al., 2016; Nagel et al., 2009). Therefore, to validate the output CityGML dataset beyond visual plausibility we have used val3dity (Ledoux, 2018), and a custom FME workbench. We tested our conversion procedures on a few datasets, and focused mostly on the one depicted in Figures 1 and 2.

A number of errors has also been included from publications of other research groups working in this domain (Section 2.1).

4. EXPLORATORY DISCOVERY OF ERRORS

Here we give specific errors that we deem are either common or particular to the IFC–CityGML conversion, without an order.

4.1 Wrong spatio-semantic paradigms

The latest proposal of CityGML 3.0 introduces the concept of constructive elements, which corresponds to the main spatio-semantic paradigm used in IFC (Figure 2). However, the currently adopted version (2.0) of CityGML does not. Nonetheless, the conversion from IFC to CityGML often results in datasets using these constructions instead of semantic surfaces, inherently resulting in invalid datasets. This inconsistency has also been discussed in Biljecki et al. (2016).

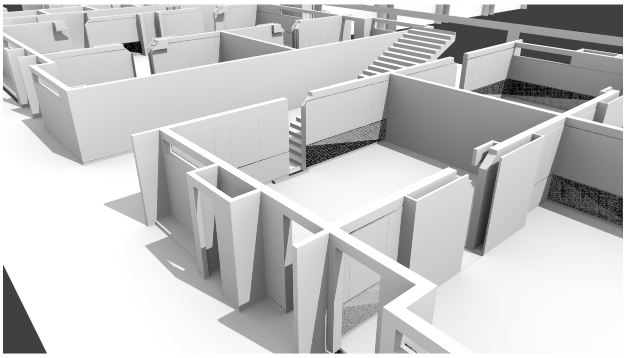

Figure 2. Walls as constructive elements instead of semantic surfaces (the same dataset as in Figure 1). These are supported in the future CityGML 3.0, but may conflict with the well ingrained semantic surface concept in 3D GIS. The image also shows overlap between geometries.

4.2 Semantic misclassification

Some feature types cannot be mapped directly between IFC and CityGML (El-Mekawy et al., 2012), due to different interpretations of semantics in the respective standards. For example, a dormer may be modelled as a building installation, or as part of the roof. Such conceptual mismatches may not be well received by software that uses the data for accomplishing a particular use case. Furthermore, such a mismatch may result in omission of features, if the mismatching types are disregarded during conversion workflow.

4.3 Omission of features

Although GIS datasets generated from BIM datasets are usually excessively detailed, the conversion from IFC may also omit features that are required for a given use case. The following reasons are possible: (i) a geometric error in the input IFC dataset prevents the conversion of a feature (Geiger et al., 2015), (ii) an IFC entity is ignored due to wrong classification (type), (iii) generalisation does not work as expected (Yu, Teo; Deng et al., 2016b), or (iv) the input dataset simply lacks semantics. Because most IFC datasets do not contain explicit distinction of interior and exterior surfaces, that information may also be missing in the conversion result. This was also noticed by Benner et al. (2005); Kang (Hong); Donkers et al. (2016); Deng et al. (2016b), who among others develop algorithms (e.g. ray-tracing) to infer exterior elements.

4.4 Commission of features

Similarly to omission, commission (excess data present in a dataset) may also occur for similar aforementioned reasons (e.g. semantic mislabelling), such as including information on the interior in lower LoD datasets.

4.5 Invalid 3D geometric primitives

In the conversion results we encountered several geometric items that do not conform to ISO19107. Some of these such as unclosed rings, wrongly orientated shells, and non-manifold solids were introduced during the conversion process. Here, the validation helped us to identify and fix errors in our algorithms.

Other errors turned out to be already present in the IFC, e.g. self-intersecting rings and shells. Figure 3 shows an example of a faulty ring. The image also shows that such errors are often tiny details in complex models. Given that native BIM authoring software exposes powerful modelling functionality that requires a lot of expertise, it can be imagined how such errors can slip in accidentally, for instance by offsetting a single face.

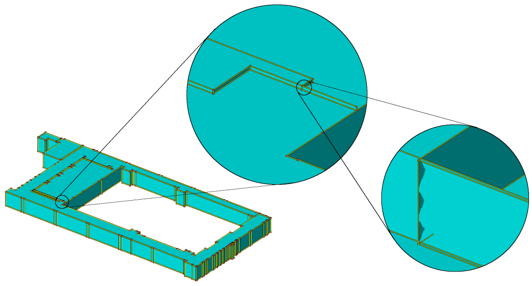

Figure 3. Self-intersecting ring in a complex geometry item, present in both original IFC and generated CityGML

Some errors could not be identified and traced to a source. They actually might be false positives caused by inappropriate sensitivity settings in the validation. For instance, we got wrong-orientation errors for very tiny wedge-shaped profiled extrusions, even though the geometry appears to be error-free on visual inspection.

Intersections in output CityGML datasets may be common due to faulty input data and/or conversion procedures (Arroyo Ohori et al., 2018), e.g. dormers overlapping with roofs (Nagel et al., 2009). Zhao et al. (2018) develop repair mechanisms to tackle these. Such errors have been much discussed in GIS, thus there is no need to elaborate on them more in this paper.

4.6 Overlapping and inconsistent spaces

Overlap and inconsistent modelling of spaces may occur often in IFC dataset (Lilis et al., 2018), potentially causing problems because CityGML does not favour overlapping features, such as rooms and building parts. Furthermore, overlapping spaces/solids may cause issues in spatial analyses, e.g. miscalculation of volume (estimating it larger than it actually is).

4.7 Lack of non-geometric information

Albeit IFC datasets are comparatively very rich, in practice they may also be coarse in terms of semantics. First, features may lack type information, preventing their mapping to semantic CityGML datasets. Second, attributes may not be populated in IFC. Third, using volumetric features (constructive elements) may interfere with semantic mapping.

4.8 Dislocated geometry

'Dislocated geometry' may be introduced in the conversion process due to faulty resolution of nested local coordinate systems in IFC to absolute coordinates in CityGML. Obvious outliers are easy to detect visually, whereas more subtle dislocations are not detectable without looking at the IFC.

4.9 High numeric values and mismatch of units

The coordinates may be a source of many errors. Owing to conversion workflows, coordinates may be stored in floating point (e.g. 2.4E9) instead of decimal representation. That is not a fault per se, but may cause errors further down the pipeline (i.e. the software that makes use of the output CityGML data does not know how to read that numeric format).

There may also be a difference between declared units and actual units of coordinates, e.g. values are in cm although the default unit is declared as meters. This can be due to the conversion either disregarding units altogether or not calculating measures properly with respect to the declared unit. Also attributes can suffer from this problem (e.g. wall width can be 300 cm but it is saved as 300 m). Errors from floating point arithmetics have also been mentioned in literature (Donkers et al., 2016).

4.10 Attribute misconversion

It is possible that some attribute information gets misconverted or wrongly interpreted. Attributes may not picked up correctly by both the conversion and utilising software due to mismatching code lists and values.

4.11 Lack of geographic references

Georeferencing of IFC data has been considered instrumental in integration with spatial data (Uggla, Horemuz). Many IFC datasets lack a spatial reference potentially resulting in difficulties. This limitation has been further described by Arroyo Ohori et al. (2017) and Barazzetti (Banfi).

4.12 Schema errors

Validating the output CityGML file against a schema is a standard procedure regardless of the acquisition technique. However, in this context schema validation is relevant to highlight because of the frequent use of the CityGML Application Domain Extension (ADE), a mechanism to extend the standard data model enabling capturing an additional set of rich information from the IFC source, potentially benefiting applications. ADEs go hand in hand in developing a conversion workflow, and in fact their use in this area has been documented in the literature (de Laat, van Berlo; Sebastian et al., 2013; Biljecki et al., 2018). Using a more complex data model such as ADE may aggravate schema errors. In our project, in which we have developed an ADE, validating the output against the ADE was a principal task, and detecting inconsistencies related to the ADE helped us to improve the workflow.

4.13 Presence of solids with interior voids

Solids with interior voids are nominally supported by both IFC and CityGML. However, they may be difficult to handle, and conversion procedures and software may simply discard them.

4.14 Redundancy of geometry

Due to the increased amount of geometry, it is important to reduce redundancy, such as avoiding storing the same geometry more than once. While not reusing equal geometries is not an error, it may prohibitively increase the complexity of data. For this purpose, CityGML enables the use of xlinks. However, they are not always used in practice (Biljecki et al., 2015).

4.15 Target software shortcomings

The software at the target (e.g. software to use the output CityGML data to estimate the noise pollution) may not be advanced enough to handle an enriched schema or excess amount of information. For example, it may not support ADEs, it may not be resolving xlinks, and during the visualisation some features may not be portrayed correctly (e.g. due to a faulty geometry engine).

5. DISCUSSION AND QUALITY FRAMEWORK

5.1 Taxonomy of errors

While different taxonomies of errors have been developed in the field (e.g. grouping them into semantic, geometric, and schematic errors), we focus on two categorisations: by source of error and by impact in applications.

5.1.1 Source of errors The three primary sources of errors we identify are the input (IFC[1]) dataset, the conversion implementation, and the software on the other end making use of the output data (i.e. implementations of use cases and visualisation).

Input-induced errors. Many errors are caused by imperfect IFC input and they propagate to the CityGML output through the conversion. Such errors are caused independently of the conversion, and would likely occur even in a properly developed conversion procedure. Some of these errors stem from native BIM or the export to IFC (Jeong et al., 2009).

Conversion-induced errors. A subset of errors are induced by the shortcomings of the conversion procedure/implementation. This exploratory work was useful because it helped us to improve our conversion implementation and developed concepts.

Utilisation-induced errors. After the conversion, the software that is reading and utilising the output dataset (CityGML in our case) may not be suited to deal with the datasets sourced with the conversion from another format. As generic as it sounds, this source of error may be considered to a large extent specific to the context of the BIM–GIS integration because of some quirky circumstances, e.g. inability to deal with the very high LoD specific to architectural datasets.

We deem these three categories almost if not entirely exclusive, e.g. we have not encountered errors that occurred due to a peculiar combination of an input dataset and conversion procedure, resulting in another type of error. However, our experience is limited to our project and a few discussions with other research groups, thus we do not exclude that this scenario may arise in a more comprehensive investigation.

[1] The terms input and output datasets are used in a generic manner that can be applied to any other data model in the BIM–GIS integration world, i.e. in the conversion from CityGML to IFC the roles would be reversed.

5.1.2 Impact of errors

Usually, conversion is driven by a particular use case (see Section 2.1). That is why it is important to consider applications when discussing errors, as the same issue may entirely prevent a use case or may not have any impact in a different scenario. Technically the latter would not be an error, but for cross-use-case analysis we call the error *benign* if it has not impact, and *malign* otherwise.

5.2 2D error matrix and examples

The combination of these two error classifications results in a matrix we would like to highlight as one of the main outcomes of the research, enabling us to better understand the lineage and propagation of errors together with the eventual impact. We give examples for two use cases: (A) estimation of the solar potential of rooftops, a common application of 3D city models (Peronato et al., 2018); and (B) estimation of the volume of buildings, a spatial analysis that has use in many domains such as energy demand and population estimation, and it is known to be sensitive to topological errors (Sindram et al., 2016). The examples are given in Table 1 and Figure 4: the same error may impact different use cases in very different ways, from not posing any issue to entirely preventing a use case. Investigating how an error propagates to the output and what is its impact for different use cases opens many possibilities for further research.

5.3 Limitations of our work

The main limitation of our work is the small number of datasets we have used to gain insight into errors. Obviously there are more types of errors than we have caught, and they highly depend on different datasets and the design of the conversion procedure. We have supplemented that with a literature review of papers that mention errors in this context, but also since quality was not much in their focus, we cannot guarantee that all errors have been covered. Besides having access to other IFC datasets, additional valuable insight would be possible from having access to conversion workflows of other research groups.

We have carried out this study during the transition of the community from CityGML 2.0 to 3.0. The former has been adopted several years ago, and is still the authoritative version of the standard. The latter facilitates 3D city models sourced from IFC, but is not yet passed by OGC, still subject to changes, and not adopted in software. An advantage of CityGML 3.0 is that according to the latest discussions it will provide support for constructive elements. Data that uses the new concepts will not conform to the current authoritative version of the standard.

We have limited the validation to the final CityGML output data, tracing sources of errors to IFC only through ad-hoc inspection instead of rigid validation. While discussing the final CityGML models we have not much focused on the quality of the information transfer process. Finally, the early nature of this work and the form of a short conference paper do not allow us to define a comprehensive framework with strict terminology, potentially rendering some of the concepts subjective.

5.4 Directions for establishing a complete framework

For future work we plan to work on a comprehensive framework supported by experiments on use cases. A priority would be to identify different sources for errors with a higher granularity, consequently we would need to take into account the whole process of generating the model. We can then identify different data generation stages (e.g. design model, native BIM, and IFC) each introducing a potential source of error. Consequently to track down a source of error, we would have to look at the different intermediate data forms and validate those.

6. CONCLUSION

While validation and quality assurance of geospatial data have been covered in many research projects, not much focus has been put on assessing specifically 3D city models originating from BIM/IFC, and the topic of BIM–GIS interoperability.

In this paper we have reported our experience of implementing the conversion from IFC to CityGML and applying it to real world data. We have coupled it with experiences of other researchers, exposing frequent quality issues that are encountered in this domain. Most issues that we have encountered can be found in 3D city models sourced with other acquisition methods. However, we encountered some particular errors mostly pertaining to the input IFC data and conversion, and documented them. By presenting these findings, the contribution of this paper is conceiving a discussion in this novel topic.

We have exposed several errors that may be considered unique or characteristic to the IFC to CityGML environment, and we have developed a classification and taxonomy (category of errors and their matrix). While the topic of the conversion between the two data formats is saturated, our experience is that quality has not been much in focus, and we hope that our work will raise awareness. However, we are not surprised with this situation, because when 3D city modelling was a nascent field and the first cities started releasing their 3D data as open data, still not many researchers and practitioners had in mind quality amid the excitement of finally having large-scale 3D city models at hand. The topic of quality of 3D data followed later, akin to this paper and the topic of IFC to CityGML conversion.

A possible long-term outcome of this work and future efforts is to benchmark and assess different conversion implementations against the listed errors. Perhaps our paper will help efforts in developing a mechanism for evaluating different workflows of a BIM to GIS conversion. Furthermore, such efforts are important to consider during the development of conversion procedures, such that errors caused by the conversion can be isolated from errors caused by the IFC input dataset. In addition, we have also discussed errors that may occur during use cases.

While today (thanks to the available software and extensive research in this domain) it is not a particular challenge to detect errors, a topic that is often overlooked is the impact of these errors on the intended use of the data, rather than having a generic point of view. We deem that a quality framework should give great importance to considering applications and the impact on the intended use of data, as it can greatly vary as discussed in Section 5. Therefore we have developed an error matrix, giving more focus on distinguishing the errors in a certain use case context.

The project from which this paper stems from focuses on the conversion from IFC to CityGML, resulting in discussions mostly pertaining to that workflow. While we believe that we have covered many different topics, it would be beneficial to gain experience in the opposite direction, possibly resulting in new insights. Looking at the bigger picture, further topics for future research are formats other than IFC and CityGML, or the synchronization of integrated BIM–GIS environments. Another

Table 1. Classification of errors from the application-dependent view. Example for (A) solar potential estimation of rooftops and (B) volume computation. The first three errors are illustrated in Figure 4.

Source of error	Example of the error	Exemplary impact in use cases
1. Input dataset	(i) Mismatched semantics: a roof is misclassified	A: malign: a software implementation would likely ignore the roof because it is not classified as such. B: benign: mislabelled semantics of surfaces would likely have no impact on volume computation.
	(ii) Invalid geometry: non-planar polygon of a wall	A: benign: it should not pose a problem because the use case focuses on roofs. While walls are still used to calculate the shadow cast on nearby objects, it is unlikely that such an error would be problematic. B: malign: invalid face of a solid renders it invalid and thus unusable for volume computations.
2. Conversion	(iii) Misconversion of the value of one attribute (type of building)	A: benign: the type of a building is mostly irrelevant for solar potential estimations. B: malign: in a use case of calculating residential volume such an error would have a substantial impact.
3. Utilisation	(iv) The software is unable to read extended schema (no CityGML ADE support) containing information about the existing energy equipment in the building	A: malign (the software may ignore information that a photovoltaic panel already exist on the rooftop, thus invalidating the solar potential analysis). B: benign: in a use case of calculating residential volume such a shortcoming is irrelevant.

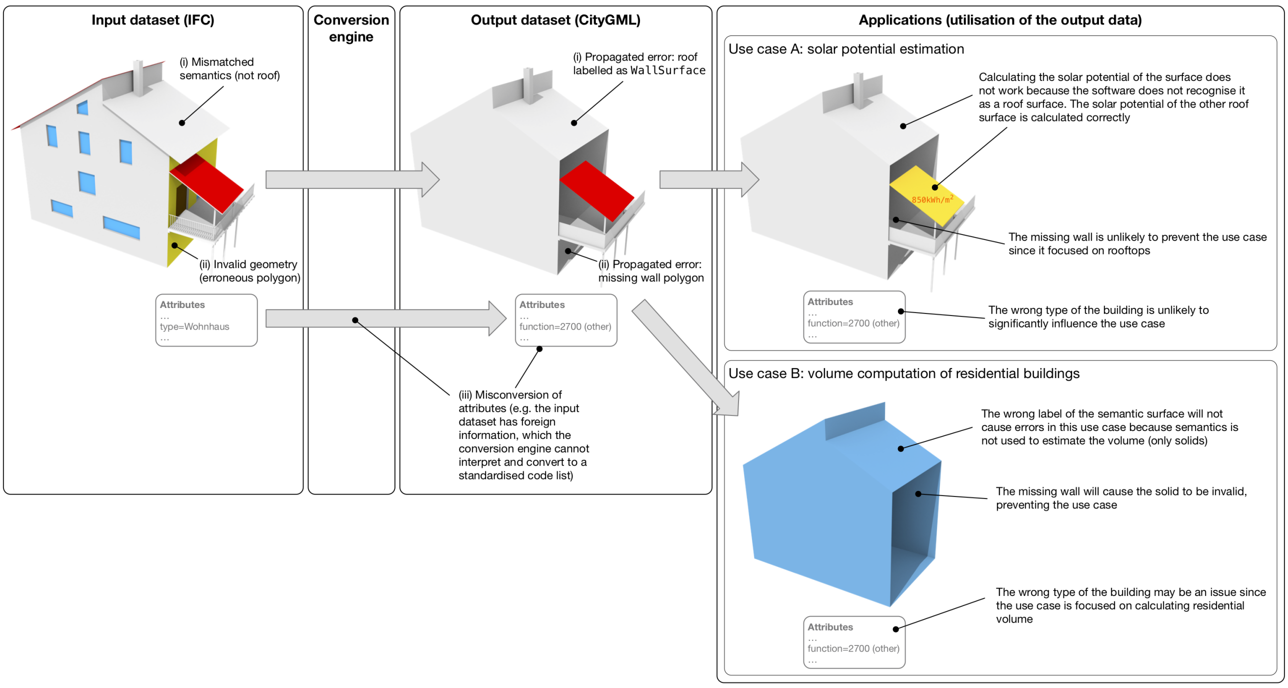

Figure 4. Combination of different types of errors results in multiple categories that are depending on the use case context. These errors are further described in Table 1. Source of the dataset (with modifications) used to generate the illustration: Institute for Applied Computer Science, Karlsruhe Institute of Technology (Häfele, 2011).

follow-up option would be a more rigid empirical experiment: investigating how simulated errors in the input IFC dataset propagate to the output CityGML dataset, and further how they take effect in different use cases.

ACKNOWLEDGEMENTS

We gratefully acknowledge the comments from the reviewers, initial discussions with Simon Lawrence (Ordnance Survey), and the IFC2CityGML project team. We thank the Building and Construction Authority (BCA) for sharing the BIM dataset. This material is based on research/work supported by the National Research Foundation under Virtual Singapore Award No. NRF2015VSG-AA3DCM001-008.

REFERENCES

Amirebrahimi, S., Rajabifard, A., Mendis, P., Ngo, T., 2015. A framework for a microscale flood damage assessment and visualization for a building using BIM–GIS integration. *International Journal of Digital Earth*, 9(4), 363–386.

Arroyo Ohori, K., Biljecki, F., Diakité, A., Krijnen, T., Ledoux, H., Stoter, J., 2017. Towards an integration of GIS and BIM data: what are the geometric and topological issues? *ISPRS*

Ann. Photogramm. Remote Sens. Spatial Inf. Sci., IV-4-W5, 1–8.

Arroyo Ohori, K., Diakité, A., Krijnen, T., Ledoux, H., Stoter, J., 2018. Processing BIM and GIS Models in Practice: Experiences and Recommendations from a GeoBIM Project in The Netherlands. *ISPRS International Journal of Geo-Information*, 7(8), 311.

Barazzetti, L., Banfi, F., 2017. BIM and GIS: when parametric modeling meets geospatial data. *ISPRS Ann. Photogramm. Remote Sens. Spatial Inf. Sci.*, IV-5/W1, 1–8.

Benner, J., Geiger, A., Leinemann, K., 2005. Flexible generation of semantic 3D building models. G. Gröger, T. H. Kolbe (eds), *Proceedings of the 1st ISPRS/EuroSDR/DGPF International workshop on Next Generation 3D City Models*, Bonn, Germany, 17–22.

Biljecki, F., Kumar, K., Nagel, C., 2018. CityGML Application Domain Extension (ADE): overview of developments. *Open Geospatial Data, Software and Standards*, 3(1), 13.

Biljecki, F., Ledoux, H., Du, X., Stoter, J., Soon, K., Khoo, V., 2016. The most common geometric and semantic errors in CityGML datasets. *ISPRS Ann. Photogramm. Remote Sens. Spatial Inf. Sci.*, IV-2/W1, 13–22.

Biljecki, F., Ledoux, H., Stoter, J., 2015. Improving the consistency of multi-LOD CityGML datasets by removing redundancy. M. Breunig, A.-D. Mulhim, E. Butwilowski, P. V. Kuper, J. Benner, K.-H. Häfele (eds), *3D Geoinformation Science*, Springer International Publishing, 1–17.

Borrmann, A., Kolbe, T. H., Donaubauer, A., Steuer, H., Jubierre, J. R., Flurl, M., 2015. Multi-Scale Geometric-Semantic Modeling of Shield Tunnels for GIS and BIM Applications. *Computer-Aided Civil and Infrastructure Engineering*, 30(4), 263–281.

buildingSMART, 2016. Industry foundation classes. Technical Report Version 4, Addendum 2, buildingSMART International Limited.

Chen, L.-C., Wu, C.-H., Shen, T.-S., Chou, C.-C., 2014. The application of geometric network models and building information models in geospatial environments for fire-fighting simulations. *Computers, Environment and Urban Systems*, 45, 1–12.

de Laat, R., van Berlo, L., 2011. Integration of BIM and GIS: The Development of the CityGML GeoBIM Extension. *Advances in 3D Geo-Information Sciences*, Springer Berlin Heidelberg, 211–225.

Deng, Y., Cheng, J. C. P., Anumba, C., 2016a. A framework for 3D traffic noise mapping using data from BIM and GIS integration. *Structure and Infrastructure Engineering*, 12(10), 1267–1280.

Deng, Y., Cheng, J. C. P., Anumba, C., 2016b. Mapping between BIM and 3D GIS in different levels of detail using schema mediation and instance comparison. *Automation in Construction*, 67, 1–21.

Devillers, R., Bédard, Y., Jeansoulin, R., Moulin, B., 2007. Towards spatial data quality information analysis tools for experts assessing the fitness for use of spatial data. *International Journal of Geographical Information Science*, 21(3), 261–282.

Donkers, S., Ledoux, H., Zhao, J., Stoter, J., 2016. Automatic conversion of IFC datasets to geometrically and semantically correct CityGML LOD3 buildings. *Transactions in GIS*, 20(4), 547–569.

El-Mekawy, M., Östman, A., Hijazi, I., 2012. An Evaluation of IFC-CityGML Unidirectional Conversion. *International Journal of Advanced Computer Science and Applications*, 3(5), 159–171.

Ellul, C., Adjrad, M., Groves, P., 2016. The impact of 3D data quality on improving GNSS performance using city models initial simulations. *ISPRS Ann. Photogramm. Remote Sens. Spatial Inf. Sci.*, IV-2-W1, 171–178.

Floros, G. S., Ellul, C., Dimopoulou, E., 2018. Investigating interoperability capabilities between IFC and CityGML LOD 4 – retaining semantic information. *Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci.*, XLII-4/W10, 33–40.

Geiger, A., Benner, J., Haefele, K. H., 2015. Generalization of 3D IFC Building Models. *3D Geoinformation Science*, Springer International Publishing, Dubai, UAE, 19–35.

Gröger, G., Plümer, L., 2012. CityGML – Interoperable semantic 3D city models. *ISPRS Journal of Photogrammetry and Remote Sensing*, 71, 12–33.

Häfele, K.-H., 2011. CityGML Model of the FJK-Haus. Institut für Angewandte Informatik (IAI), Karlsruher Institut für Technologie.

Isikdag, U., Zlatanova, S., 2009. Towards Defining a Framework for Automatic Generation of Buildings in CityGML Using Building Information Models. *3D Geo-Information Sciences*, Springer, 79–96.

ISO, 2013. Industry Foundation Classes (IFC) for data sharing in the construction and facility management industries.

Jeong, Y. S., Eastman, C. M., Sacks, R., Kaner, I., 2009. Benchmark tests for BIM data exchanges of precast concrete. *Automation in Construction*, 18(4), 469–484.

Jusuf, S. K., Mousseau, B., Godfroid, G., Soh Jin Hui, V., 2017. Integrated modeling of CityGML and IFC for city/neighborhood development for urban microclimates analysis. *Energy Procedia*, 122, 145–150.

Kang, T. W., Hong, C. H., 2015. A study on software architecture for effective BIM/GIS-based facility management data integration. *Automation in Construction*, 54, 25–38.

Kang, T. W., Hong, C. H., 2017. IFC-CityGML LOD mapping automation using multiprocessing-based screen-buffer scanning including mapping rule. *KSCE Journal of Civil Engineering*, 4(4), 1–11.

Konde, A., Tauscher, H., Biljecki, F., Crawford, J., 2018. Floor plans in CityGML. *ISPRS Ann. Photogramm. Remote Sens. Spatial Inf. Sci.*, IV-4/W6, 25–32.

Kumar, K., Labetski, A., Arroyo Ohori, K., Ledoux, H., Stoter, J., 2019. The LandInfra standard and its role in solving the BIM-GIS quagmire. *Open Geospatial Data, Software and Standards*, 4(1), 10.

Ledoux, H., 2013. On the Validation of Solids Represented with the International Standards for Geographic Information. *Computer-Aided Civil and Infrastructure Engineering*, 28(9), 693–706.

Ledoux, H., 2018. val3dity: validation of 3D GIS primitives according to the international standards. *Open Geospatial Data, Software and Standards*, 3, 1.

Lee, Y.-C., Eastman, C. M., Lee, J.-K., 2015. Validations for ensuring the interoperability of data exchange of a building information model. *Automation in Construction*, 58, 176–195.

Lilis, G. N., Giannakis, G. I., Katsigarakis, K., Costa, G., Sicilia, Á., Garcia-Fuentes, M. Á., Rovas, D. V., 2016. Simulation model generation combining IFC and CityGML data. *Proceedings of the 11th European Conference on Product and Process Modelling (ECPPM2016)*, Limassol, Cyprus, 215–222.

Lilis, G. N., Giannakis, G. I., Rovas, D. V., 2015. Detection and semi-automatic correction of geometric inaccuracies in IFC files. *14th Conference of International Building Performance Simulation Association*, Hyderabad, India, 2182–2189.

Lilis, G. N., Giannakis, G., Katsigarakis, K., Rovas, D. V., 2018. A tool for IFC building energy performance simulation suitability checking. *eWork and eBusiness in Architecture, Engineering and Construction*, CRC Press, 57–64.

Lim, J., Tauscher, H., Biljecki, F., 2019. Graph transformation rules for IFC-to-CityGML attribute conversion. *ISPRS Ann. Photogramm. Remote Sens. Spatial Inf. Sci.*

Liu, X., Wang, X., Wright, G., Cheng, J., Li, X., Liu, R., 2017. A State-of-the-Art Review on the Integration of Building Information Modeling (BIM) and Geographic Information System (GIS). *ISPRS International Journal of Geo-Information*, 6(2), 53.

Löwner, M. O., Gröger, G., Benner, J., Biljecki, F., Nagel, C., 2016. Proposal for a new LOD and multi-representation concept for CityGML. *ISPRS Ann. Photogramm. Remote Sens. Spatial Inf. Sci.*, IV-2/W1, 3–12.

Nagel, C., Stadler, A., Kolbe, T. H., 2009. Conceptual requirements for the automatic reconstruction of building information models from uninterpreted 3D models. *Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci.*, XXXVIII-3-4(C3), 46–53.

Noardo, F., Arroyo Ohori, K., Biljecki, F., Krijnen, T., Ellul, C., Harrie, L., Stoter, J., 2019. GeoBIM benchmark 2019: design and initial results. *Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci.*, XLII-2-W13, 1339–1346.

OGC, 2016. OGC CityGML Quality Interoperability Experiment. Technical Report OGC 16-064.

Open Geospatial Consortium, 2012. OGC City Geography Markup Language (CityGML) Encoding Standard 2.0.0. Technical report.

Peronato, G., Rey, E., Andersen, M., 2018. 3D model discretization in assessing urban solar potential: the effect of grid spacing on predicted solar irradiation. *Solar Energy*, 176, 334–349.

Rafiee, A., Dias, E., Fruijtier, S., Scholten, H., 2014. From BIM to Geo-analysis: View Coverage and Shadow Analysis by BIM/GIS Integration. *Procedia Environmental Sciences*, 22, 397–402.

Sargent, I., Harding, J., Freeman, M., 2007. Data quality in 3D: gauging quality measures from users' requirements. *Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci.*, XXXVI-2/C43, 8.

Sebastian, R., Böhms, M., van den Helm, P., 2013. BIM and GIS for low-disturbance construction. N. Dawood, M. Kassem (eds), *Proceedings of the 13th International Conference on Construction Applications of Virtual Reality*, London, UK, 469–479.

Sindram, M., Machl, T., Steuer, H., Pültz, M., Kolbe, T. H., 2016. Voluminator 2.0 – speeding up the approximation of the volume of defective 3D building models. *ISPRS Ann. Photogramm. Remote Sens. Spatial Inf. Sci.*, III-2, 29–36.

Stouffs, R., Tauscher, H., Biljecki, F., 2018. Achieving Complete and Near-Lossless Conversion from IFC to CityGML. *ISPRS International Journal of Geo-Information*, 7(9), 355.

Sun, J., Harrie, L., Jensen, A., Eriksson, H., Tarandi, V., Uggla, G., 2018. Description of geodata quality with focus on integration of BIM-data and geodata. Technical report, Stockholm, Sweden.

Tauscher, H., Stouffs, R., 2019. Extracting different spatiosemantic structures from IFC using a triple graph grammar. *Intelligent and informed: 24th Annual Conference of the Association for Computer-Aided Architectural Design Research in Asia (CAADRIA 2019)*, Wellington, New Zealand, 605–614.

Uggla, G., Horemuz, M., 2018. Geographic capabilities and limitations of Industry Foundation Classes. *Automation in Construction*, 96, 554–566.

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.

Wang, H., Pan, Y., Luo, X., 2019. Integration of BIM and GIS in sustainable built environment: A review and bibliometric analysis. *Automation in Construction*, 103, 41–52.

Yu, S.-C., Teo, T.-A., 2014. The generalization of BIM/IFC model for multi-scale 3D GIS/CityGML models. *Proceedings of the 35th Asian Conference on Remote Sensing*, Nay Pyi Taw, Myanmar.

Zhao, J., Ledoux, H., Stoter, J., Feng, T., 2018. HSW: Heuristic Shrink-wrapping for automatically repairing solid-based CityGML LOD2 building models. *ISPRS Journal of Photogrammetry and Remote Sensing*, 146, 289–304.

Zhao, L., Liu, Z., Mbachu, J., 2019. Highway Alignment Optimization: An Integrated BIM and GIS Approach. *ISPRS International Journal of Geo-Information*, 8(4), 172.