A TESTBED FOR APPLYING OPERATIONALIZED GRAPH GRAMMARS TO SUSTAINABILITY ANALYSIS IN INTEGRATED BIM-GIS SCENARIOS

H. Tauscher¹, K.W. Wong²

¹ Bauhaus-Universität Weimar, Germany - helga.tauscher@uni-weimar.de
² HTW Dresden, University of Applied Sciences, Dresden, Germany - kawing.wong@htw-dresden.de

Commission IV, WG IV/9

KEY WORDS: IFC, CityGML, BIM, LCA, Thermal Energy Simulation, Solar Analysis, Graph Transformation

ABSTRACT:

Environmental aspects are becoming increasingly important for construction projects and thus holistic analysis spanning different scopes such as the building and city context. With graph-based data integration we propose a general method to support cross-domain analysis tasks. In this paper, we present a testbed consisting of a data set with three university buildings in IFC and CityGML format and two use cases for sustainability analysis spanning the building and city scope. We discuss the details of the data set and the two use cases - solar irradiation and material impact analysis - and show how graph-based data integration will be carried out in the testbed.

1. INTRODUCTION

1.1 Motivation and objective

In architectural design and construction planning, the past decades have seen environmental aspects such as energy-efficiency, resource preservation and ecosystems protection being added as criteria and principles to the classic Vitruvian triad of firmitas (stability), utilitas (utility) and venustas (beauty). Thus, design and planning have increasingly included holistic views and the respective analysis transcends the scope of a single building in its completed state both in space and time. Life-cycle assessment (LCA) is a holistic endeavour that does not only cover the various temporal stages of planning, erecting and using a building but also the different spatial scopes from the interior facilities to the surrounding built and natural environment. Consequently, there is a necessity to integrate semantic 3D models of the building and the city scope, namely building data in industry foundation classes (IFC) format and city data in City Geography Markup Language (CityGML) format.

There have been various attempts during recent years, to tackle integration between these types of data. Yet, most proposals are still subject to certain limitations. They focus on singular operational integration cases (e.g. unidirectional transformation). They are hard-coded and inflexible at runtime. They are not suitable for domain-specialists to handle, and they don’t allow for formal verification or quantitative assessment. In previous work, we have proposed an integration approach where operational cases such bidirectional transformation and synchronization are derived from generic rules (Tauscher 2020). Beyond generic support for the various operational cases, this approach is also based on a runtime-interpreted declarative domain specific language with an intuitive visual representation suitable for domain specialists and based on a rigid formal framework.

In this paper, we are presenting a test bed for the generic integration approach to demonstrate applicability of the method to the area of sustainability assessment. In Section 2 we describe the test bed as such, consisting of a two-part data set and two use cases for sustainability analysis. In Section 3 we then discuss the data integration necessary for the use-case-specific analysis tasks to be carried out, show how it would be tackled using traditional engineering methods and how graph based methods will subsequently applied to the testbed.

1.2 Related work

Most sustainability analysis for the built environment is to be carried out across the building and city context. (Wang et al. 2019) conducted a review of 76 publications related to BIM-GIS integration for the purpose of creating a sustainable built environment. From the research literature, they identify the most prominent aspects of BIM-GIS integration: highly integrated visualization, two-way interactive data flows, open standards and specifications, customization, and user-friendly experiences. In conclusion they propose that GIS should be involved more in construction projects. We have selected two aspects of sustainability that are already subject of extensive analysing on both the BIM and GIS level and scale, (a) solar and thermal energy and (b) life-cycle impact of incorporated materials. For both cases, we summarize the state of the art in the following.

Solar potential evaluation using GIS is considered mature, particularly for rooftop areas. The commercial software ArcGIS, for example, contains a tool set consisting of 2D analysis to calculate irradiation with the viewshed method as well as 3D analysis to calculate shadow volumes and frequency. Accumulated irradiation and shadow volumes can be combined to obtain irradiation values per defined surface and time slot. There is open source software with similar functionality, such as the QGIS plug-in UMEP with a processor for simulating solar irradiation on whole building envelopes (Lindberg et al. 2019).

Footnotes:

* Corresponding author
the Blender add-on VI-Suite employing the raytracing software Radiance to calculate irradiation [Southall and Biljecki (2017)].

Thermal energy simulation on the level of single buildings is usually employed in planning processes to calculate thermal gain under design and technical installation variants. These simulations require input obtained from the GIS context. For example, EnergyPlus expects net thermal radiation heat gain as rate [W], rate per area [W/m²] or Energy [J] as input for a particular surface [EnergyPlus Development Team (2022a,b)]. In addition shading surfaces can be given to be considered in solar gain calculation where the beam (direct) radiation proportion is attenuated with a factor of the sunlit surface proportion. For these calculations, shadow-casting surfaces of the building’s environment have to be remodelled in EnergyPlus. There are further uses of solar irradiation values in design and planning processes of new building or refurbishment projects such as urban farming site identification [Palliwal et al. (2021)].

Analysis of material incorporated in a construction accounts for the life-cycle emissions of construction materials. Digital building models are able to contribute the required information for carrying out these analysis and there are methods for efficient retrieval of bulky data, even for buildings made of complicated materials or large-scale construction, as shown by Veselka et al. (2020) and Teng and Pan (2018). The available methods to calculate these values range from fully automated workflows to manual or semi-automatic determination from Bill of Quantity data exported from CAD software with standard functionality.

While material analysis on the BIM level is mature, larger scale analysis of material stock and flows on the GIS level appear as an aspiring research area. Current approaches estimate material stock on the base of building type information from cadastral sources [Hartmann et al. (2016)] or with semantic enhancement from imagery [Haberl et al. (2021)] and subsequent assignment of empirically derived material intensities (ibid.) or material composition indicators [Ortlipp et al. (2018)] per building type instead of the results of actual material analysis or survey.

2. TESTBED

The data set covers three university buildings on the campus of HTW Dresden — University of Applied Sciences in Germany both as detailed building models and as contextualized city models. In Section 2.1 we describe where this data set is acquired from, how it was examined, checked, enhanced and improved. Besides data, the testbed contains two exemplary use cases for cross-domain analysis. In each use case, one part of the analysis is carried out on one domain model and the other part on the domain model dataset. Hence, we have an holistic analysis that crosses the boundaries of domain models. In the following sections, we describe the particular methods selected for implementing the use cases in our test bed: The first case is an analysis of the irradiation potential of building surfaces (Section 2.2). The second case is a life-cycle assessment (LCA) of the analysed building materials used in buildings (Section 2.3).

2.1 IFC and CityGML data

The data set covers three university buildings on the campus of HTW Dresden: the main building (Z-building, 1962), one seminar building (S-building, 1955) and a laboratory building (N-building, 2002), see Figure 1. Given the representation of different construction periods and academic functions, the ensemble of the three campus buildings constitutes an interesting target for environmental analysis.

The building models in IFC format originate from earlier student projects. They are authored with Nemetschek Allplan (version 2021). The common elements of the three models with directly measurable quantities, size, shape, location, and orientation, can be approximately classified as “Level of Development (LOD) 300” with reference to BIMForum (2021). Two additional enhancements on the models were done in the native CAD application before the export to IFC. Firstly, “Room” elements, which represent indoor space, were added for net floor area (NFA) calculation. Another critical step is to ensure the complete attachment of the material information as the “Material” attribute to all relevant building elements. Material information was not included in the models and had to be obtained from historical engineering drawings and documents. It is noticeable that extra attention should be paid for the building elements with multiple materials. Using layered walls as example, material information should be assigned to the corresponding wall layer (IFC entity type IfcMaterialLayer) instead of the whole wall (IFC entity type IfcWall), so as to prevent overvalued quantities assigned to particular materials. When correct material information can be found in the material attribute of building element in the exported IFC file, the material information is well stored and the file is ready to be extracted for LCA calculation. The exported IFC can also be enhanced by auxiliary software such as IFC Optimizer.

The city model containing the three buildings and their surrounding buildings (a segment of 2 km x 2 km) in CityGML format was downloaded from the open geodata portal provided by the Federal State of Saxony and the municipal surveying authority [3]. The CityGML file in Level of Detail (LoD) 2 was

3 Staatsbetrieb Geobasisinformation und Vermessung Sachsen, GeoSN: https://www.landesvermessung.sachsen.de/digitale-hoehen-und-
imported to ArcGIS Pro (version 2.6.1) for preprocessing and later analysis. During import, the geometries are converted into the internal multipatch package format (.gdb) which contains 3D polygon layers of wall and roof surfaces of respective buildings. After visually checking the data such as dimensions of the 3D polygons, the city model can be further processed in response to the simulation design. In particular, in order to obtain a more precise result for the distribution of irradiation on various rooftop sections with respect to their height and orientation, roof polygons of the three buildings were processed with merging and splitting tools and by adjusting their heights. In addition to the 3D city model, the digital surface model (DSM) for the same section of 2 km x 2 km has been obtained from the aforementioned source.

### 2.2 Irradiation and shadow analysis

The first use case is about the irradiation potential of building roof surfaces. In this analysis, we assess the cross-shadowing between several buildings on the city scope and then use the results for thermal energy simulation on the building scope. Figure 2 shows the conceptual flow of analysis with the city model on the left and building model on the right.

The irradiation potential assessment consists of two parts. First, the direct simulation of solar irradiation using the "Area Solar Radiation" geo-processing tool in ArcGIS with CityGML and DSM as the main inputs. Other input parameters such as the diffuse lighting proportion were customised to the conditions in Dresden. After the simulation, the major output is the annual irradiation in $Wh/m^2$. Total irradiation in $MWh/year$ in each roof section can be calculated by multiplying the mean irradiation with the surface area. To further rate the irradiation potential, these values where further multiplied with the mono-crystalline photovoltaic modules' efficiency (18%) and installation performance ratio (85%) as quoted by [Philips and Warmuth (2022)](https://doi.org/10.5194/isprs-archives-XLVIII-4-W4-2022-147-2022), resulting in the potential power generation by solar panels in $kWh/year$.

Second, shading simulation has been carried out in ArcGIS with the "Sun Shadow Volume" and "Sun Shadow Frequency" tools based on the 3D polygons from CityGML file as the main input feature. The outcomes are shadow volume models (in the form of polyhedrons, see Figure 3) as well as shadow frequency (in the form of shading times per cell in raster layer, see Figure 4) for specific date and time slots. For the performed analysis, sunrise and sunset times on dates of summer solstice, winter solstice, spring equinox and autumn equinox of Dresden were chosen to represent the longest, shortest and medium day-lengths, leading to analysis results for extreme and general scenarios.

With the quantitative solar potential in the form of attributes and features representing irradiation, shadow frequency and shadow volume generated from the GIS domain, the city model can be enhanced and after transferal to corresponding roof or slab entities in the building model as new properties and entities in IFC, the solar potential can be further assessed with the methods from within the BIM domain to be used in solar heat gain calculation, thermal energy simulation or plannings on renewable energy installation, vegetation and plantation.

### 2.3 Life-cycle assessment of building materials

The second use case is about the life-cycle assessment of materials used in buildings. In this analysis, we assess the material quantities and their environmental impact on the building scope and then use the accumulated results for large-scale (visual or algorithmic) analysis of the LCA key values on district, city, regional, national or even global level. Figure 5 shows the conceptual flow of analysis with the building model on the right and the city model on the left.
LCA is carried out according to the requirements and guidelines of the certification scheme issued by the German Sustainable Building Council (DGfB) with LCA being one of 6 certification criteria (ENV 1.4). We limit ourselves to the simplest calculation option, the Partial Calculation Method (PCM). According to DGfB guideline, when applying the PCM, only materials used by major and load-bearing building elements are accounted for. Further, we restrict the use case to effects embodied in construction materials and ignore those resulting from the operation of technical equipment and machines during the various lifecycle phases.

The quantities of the building parts required according to the PCM together with the corresponding materials of the three campus buildings were extracted from the building models. Native CAD software provides methods to export bills of quantities, for example in spreadsheet or printed report form, but we prefer to operate on application-independent data in IFC format. Most IFC-capable software provides functionalities to calculate quantities and export those for further processing. Software that is scriptable or extendable, like FreeCAD or the Opensource BIM server, can also be used to implement calculations following below. For setting up the testbed, we have, however, resorted to working with a spreadsheet.

First of all, materials from the recognized database (OKOBAU) to working with a spreadsheet. For setting up the testbed, we have, however, resorted to effects embodied in construction materials and ignore those resulting from the operation of technical equipment and machines during the various lifecycle phases.

The quantities of the building parts required according to the PCM together with the corresponding materials of the three campus buildings were extracted from the building models. Native CAD software provides methods to export bills of quantities, for example in spreadsheet or printed report form, but we prefer to operate on application-independent data in IFC format. Most IFC-capable software provides functionalities to calculate quantities and export those for further processing. Software that is scriptable or extendable, like FreeCAD or the Opensource BIM server, can also be used to implement calculations following below. For setting up the testbed, we have, however, resorted to working with a spreadsheet.

First of all, materials from the recognized database (OKOBAU) could be obtained from the corresponding material available in the IFC file. In a second step, LCA factors of the materials could be obtained from the IFC file. In a second step, LCA factors of the materials could be obtained from the corresponding data in the database. Third, and finally, calculations following the PCM guidance were done on Excel spreadsheet, resulting in the final life-cycle environmental impact emissions, in terms of global warming potential, ozone depletion potential, photo-chemical ozone creation potential, acidification potential and eutrophication potential of each individual building.

Given a set of structural components $C$, a set of materials $M$, and a set of environmental impact indicators $I$, the emissions are calculated with equations

$$E_{i,c} = \sum_{m \in M} (Q_{i,m} \cdot F_i)$$

$$E_{i,m} = \sum_{c \in C} (Q_{c,m} \cdot F_i)$$

$$E_i = \sum_{c \in C} \frac{1}{TA} \sum_{m \in M} \frac{1}{TA} E_{i,m}$$

where

$E_{i,c}$: Emission of a particular environmental impact indicator $i \in I$ of a particular main structural component $c \in C$, across all materials, in [kg] emission equiv.

$E_{i,m}$: Emission of a particular environmental impact indicator $i \in I$ of a particular material $m \in M$, across all components, in [kg] emission equiv.

$E_i$: Summation or total emission of a particular environmental impact indicator $i \in I$ of a whole building, across all components and materials, in [kg/(m²·a)] emission equiv.

$Q_c$: Quantity (based on assigned functional unit in database) of a particular main structural component $c$, in [m²] or [m³], [m] or [kg]

$F_i$: Emission factor of environmental impact indicator $i \in I$ summed over the life-cycle phases: A1-3 (production), C3 (waste processing), C4 (disposal) and D (recycling or recovery potential), $F_i = F_{A1-3} + F_{C3} + F_{C4} + F_D$, in [kg/m²] or [kg/m³] or [kg/m] or [kg/kg]

$T$: Intended duration of use of the evaluated building, with reference period of 50 years, in years [a]

$A$: Net floor area, in [m²]

The embodied carbon of each building, which is indicated by the calculated life-cycle global warming potential, in form of total carbon emissions in [kg/(m²·a)] $\text{CO}_2$ equiv., will be imported to the city model being attached as new attributes of the three campus buildings. The intensity of embodied carbon can then be visualized by different colors, providing a macroscopic view for further analysis regarding to district planning.

### 3. DATA INTEGRATION

In this Section, we develop the necessary integration between the two domain models for the use cases described in Sections 2.2 and 2.3. Figure 5 demonstrates how it would be tackled using traditional engineering methods (Section 3.1) or how graph based methods will be operationalized for application to the testbed with these use cases (Section 3.2 and further use cases (Section 3.3).

#### 3.1 Correspondence graph: Manual method

In the course of setting up the testbed, we carry out the necessary integration manually. As a precondition for the holistic analysis, we have to establish a connection between the building and city model data sets. To this end, we create a table of links between the city model section and the three building models represented as comma-separated values (CSV format), expressing semantic equivalence relations between IDs of elements relevant for the analysis in both domains, as shown in Table 1. This approach with an independent link model in addition to the domain models corresponds to the multi model concept introduced by Fuchs and Scherer (2017).

We take into account the problems and issues with semantic equivalence described by Beck et al. (2021), but still consider some sort of well-defined equivalence as the only viable way to bridge the domains. We are using the following elements appearing on both sides for creating equivalence relations: a single building as such (for use case 2) and the surfaces

---

Figure 5. Use case 2: Material life-cycle assessment on the building model (right) followed by large scale material stock analysis on the city model (left)
of the outer building shell (for use case 1). While the correspondences between buildings can be established in a straightforward manner for this particular project, subtle differences in how buildings are identified and separated in IFC and CityGML, for example at the conjunction of Z- and S-building with an interconnecting building part above a passage, may hinder the attempt. This is even more evident when looking at surface correspondences. Surfaces are involved in building models as bounding entities of building elements. In theory, correspondences can be established between pairs of surfaces on the IFC and CityGML side and for pairs of building elements, because schemata correspond in both concepts (surfaces and elements). In practice, links will have to be between CityGML surfaces and IFC building elements, because the respective other concept is not instantiated in commonly found models.

The manual linking process can be supported by querying methods and leverage similarity of geometric (e.g. spatial proximity), semantic (e.g. external location in building shell), as well as structural (e.g. bounding of corresponding spaces) properties of the linked elements.

After correspondences are established, between analysis steps, when moving from one domain to the other, we are updating attribute values in the respective other model following along these connections as is illustrated with Figure 6.

### 3.2 Correspondence graph: Automated method

The testbed is designed for testing the various operational scenarios of graph-based integration. One of these operational data integration scenarios is correspondence generation. In this scenario, the corresponding surfaces, that have been manually identified before, would be identified automatically. Even when supported by advanced querying methods in the domain models, manual correspondence generation remains tedious work. Description of the data integration via graph transformation will incorporate the queries from the manual process into productions or rules, such that the process of correspondence generation is reproducible and thus can be automated. Similar to the manual linking, automated matching can be based on spatial structure, semantic features and geometry.

Another, more extensively investigated operational data integration scenario is conversion, mainly in the direction from the building to the city model, see for example the IFC2CityGML project. Conversion in this direction can be employed to update city models from planning data. It can also be used for the enrichment of urban environment surveying results with results of downstream survey of interior buildings. The graph transformation systems derived for the conversion would generate the required links on the fly, independent of the targeted direction. Finally, the creation of sample data with graph transformation would directly operationalize the generic transformation system in the mode of model creation and would similarly generate correspondences together with the domain models.

### 3.3 Synchronization for testbed use cases

For carrying out the use cases described in Section 2 we operationalize building and city model integration in the mode of synchronization. We keep independent building and city models with their pre-generated correspondences and propagate changes from one domain side to the respective other domain between analysis steps using forward and backward synchronization. This way we avoid the creation of redundant data. This would be inevitable with an alternative approach of converting all necessary information to either the city model or the building model in order to carry out the whole analysis with multiple steps in one single model. For use case 1 this would mean to transfer show-casting surfaces from the city to the building model and for use case 2 this would mean to transfer building details on a large scale from the building to the city model. With separate models, we are also able to leverage specific tools for parts of the analysis which may only be able to process data from one of the two domains.

### 4. CONCLUSION

#### 4.1 Summary

In this paper we presented a testbed for graph-based data integration methods (Section 2) and sketched the application of the various operational methods of graph-based data integration in the context of that testbed (Section 3.3). The testbed contains a data set and use cases suitable for verification of the advised methods.

#### 4.2 Limitations

For use case 2, the LCA has been confined to materials and excluded environmental impact not embodied in construction materials. It is well known that a considerable amount of energy consumption is caused in other ways during the use and operation of a building. Although some of the missing LCA factors can be supported with information from digital building...
models as well, the material analysis seems to be a sufficient and appropriate portion to study the necessary data integration.

For use case 1, the elements and options of irradiation analysis have been identified and demonstrated, but there are still gaps to close, in particular with regard to vertical surfaces and specification of the result to be transferred. Also, we have employed commercial software for the analysis. As the testbed aims to serve reproducibility of the research, it would be consequent to change over to open source software, such that future studies are easily reproducible, given that conversion procedures and rule sets would also be openly available.

Links have been created only exemplarily and need to be completed.

4.3 Outlook

In this paper, we presented a testbed for graph-based data integration of building and city data, but the actual application of graph transformation and a respective rule set is still under development. Later, when using the testbed to prove the concept and power of graph grammars for data integration, we aim to show that the same rule set can be used for different operational cases. We will not only study the synchronization case, but also forward and backward conversion as well as link generation. With link generation by graph transformation we can automate the preliminary step of identifying equivalent elements on both domain sides. This corresponds to the concept of explicit and implicit link specification formulated by Fuchs and Scherer (2017).

With the help of this data set and use cases we can also answer more technical and broader questions of applicability of graph transformation to the built environment, e.g. impact of model sizes on performance, sparse reading of larger models, optimal splitting and distribution of models and partial updates in design and planning process. The testbed can further serve to reproduce and compare other approaches to data integration that are not graph-based or employ graphs in a different way.

For the application described in Section 3.3 we need to extend the existing rule set and evaluate it with Emoflon. This will be possible with the given data. More work is needed for the application described in Section 3.2. In particular, rules involving geometric constraints would possibly be necessary. We hope that the testbed will aid the substantial research still required towards this goal.

ACKNOWLEDGEMENTS

The authors would like to thank Prof. Undine Kunze (Chair of Construction Informatics, HTW Dresden) for generously providing access to sample data of the university buildings for testing purposes and Heike Schuster (Technical department, HTW Dresden) for providing access to technical documents and advise on their interpretation.

References


Teng, Y., Pan, W., 2018. BIM-integrated embodied carbon assessment of prefabricated buildings in Hong Kong. International Conference on Construction Futures.
