ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences, Volume X-4/W7-2025
9th International Conference on Smart Data and Smart Cities (SDSC), 2–5 September 2025, Kashiwa, Japan

Structuring Subsurface Knowledge: Data Management and a Modeling Framework for an Integrated Geological and Urban 3D Model — A Case Study for the City Center of Stuttgart

Klaus Pusacker [1,2], Volker Coors [2], Isabel Rupf [1,2] , Matthias Ebert [1], Sisi Zlatanova [3], Ben Gorte [3]

[1]Regierungspräsidium Freiburg - Dept. 9: State Authority for Geology, Mineral Resources an Mining (LGRB), Albertstraße 5, 79104 Freiburg i. Br., Germany - (Klaus.Pusacker, Isabel.Rupf, Matthias.Ebert)@rpf.bwl.de
[2]Stuttgart University of Applied Sciences, Schellingstr. 24, 70174 Stuttgart, Germany - Volker.Coors@hft-stuttgart.de
[3]University of New South Wales, Faculty of the Built Environment, Kensington Campus, Sydney Australia, (S.Zlatanova, B.Gorte)@unsw.edu.au

Keywords: Geological 3D modeling; Urban underground modeling; Voxels; CityGML; Case study Stuttgart

Abstract

In subsurface planning of urban areas, the use, management, and exchange of urban and geological data serve as a fundamental basis for collaboration among various specialist in the fields of urban and geological 3D modeling. Therefore, the development of a data model schema tailored to urban and geological subsurface models is an important foundation for data management. However, various information's about buildings, infrastructures, and geological structures, that cannot be integrated in common data model structures requires an expansion and combination of the existing data model structures for geological and urban 3D models. The existing data model schemas can already manage various types of subsurface data. Building on this, this paper aims to advance the development of a data model by combining existing schemas to meet the requirements for data management in 3D voxel modeling of the subsurface. The conception of such a data model is implemented for a case study in the City of Stuttgart, based on the GML application schemas Geoscience Markup Language (GeoSciML) and City Geography Markup Language (CityGML). Our data model is applied using urban and geological model elements for the Stuttgart case study. In this context, the construction and integration of the model elements within a framework that comprises modeling, data management, and visualization tools were examined.

1. Introduction

In urban subsurface modeling, various professionals from the fields of 3D city modeling and geological 3D modeling, collaborate to address different tasks with varying requirements for modeling data (e.g. Coors et al., 2022; Otón et al., 2021). In this context, large volumes of urban planning and geological data concerning buildings, infrastructures, and geoscientific structures can be compiled, analyzed, and evaluated.

In geotechnical investigations, large amounts of data are collected from a wide range of geological, geotechnical, and hydrogeological measurements of subsurface layers. Using modern methods for evaluating commonly used geoscientific data sources such as boreholes (Zhang et al., 2023), profile sections (Niu et al., 2024), and geological maps (Ran et al., 2022), the structure of the subsurface is interpreted and reconstructed in plans or models.

For urban planning and the development of city models, the interoperable integration, management, and use of large datasets from various sources are increasingly being enabled. Virtanen et al. (2024) describe the integration of remote sensing and other sensor data. Kasprzyk et al. (2024) have examined the implementation of the interoperable CityGML standard for various database systems that are suitable for managing large datasets. Methods for using artificial intelligence in city modeling are also being developed as investigated by Komar and James (2024).

Many geological and geotechnical subsurface information sources are evaluated in an application-specific manner. This leads to the loss or simplification of irrelevant information for the application due to the lack of data structures in which all subsurface information can be managed for interdisciplinary use (Wu et al., 2021; Tegtmeier et al., 2014). Jeong et al. (2024) emphasize that in the broad scope of Smart City projects, data and their relationships are often not understood by non-expert users. Therefore, a data model adapted for urban and geological subsurface modeling should be able to capture a wide range of subsurface data. To optimize data exchange and the shared use of information, subsurface data must be archived in various formats and in a harmonized manner.

The present study focuses on the construction of model elements and their management within a newly designed data model. Moreover this study aims to achieve two main goals. The first is to reduce data loss by developing a new data model that not only represents the model elements themselves but also references the underlying base data. The second is to present the efficient octree data structure designed for voxel-based models. Accordingly, this study focuses on the following research questions:

1. How do model elements need to be constructed and processed to enable their integration into a unified geological and urban 3D model?

2. What voxel resolution is required for the accurate representation of different model elements?

3. What resolution can a octree data structure support and how much storage space is necessary for data management?

2. State of the Art

2.1 Data Models for Urban Subsurface Modeling

To ensure compatibility and relevance, widely used data models in urban and geological information management were assessed in advance of this study. In the field of geological data management, GeoSciML (CGI, 2021) has been established as an application schema, while in the field of urban and building modeling, the data model schemas CityGML (Kolbe, 2021) and the Industry Foundation Classes (IFC) (Building Smart, 2022) are well known.

GeoSciML is a data model and encoding standard with a focus on "interpreted geology," as it is visualized on geological maps or in 3D models (CGI, 2021). GeoSciML defines a clear structure of classes in which both information on common geological strata and structural elements, as well as diverse data from geoscientific measurements, can be managed (Qu et al., 2024; Tegtmeier et al., 2014).

IFC is an open-source data exchange format that has become the standard for Building Information Modeling (BIM) within the construction and facility management sectors. As outlined by (Chapman et al., 2022) , BIM serves as both a methodology for interdisciplinary planning and process coordination, and as an interoperable modeling approach widely adopted in building and facility management in Stuttgart. For a comprehensive overview of the technical foundations of CityGML and its applications in 3D city modeling and urban planning in Stuttgart, see also Kolbe et al. (2021) and Padsala et al. (2021).

The integration of GeoSciML with CityGML and IFC has been explored in a limited number of studies as a potential foundation for developing data models for urban subsurface modeling. Köbberich et al. (2022) assigned a customized extension of the GeoSciML data model within an IFC data model to integrate geological and geotechnical data into building models according to the BIM standard. A conversion of GeoSciML data into the IFC format was designed to provide geoscientific data for building and infrastructure models on a project-specific basis. For this purpose, converters were also developed that allow users to convert their borehole profiles, voxel datasets as well as geological boundary surfaces and envelopes into IFC datasets. The 3D GEM data model by Tegtmeier et al. (2024) harmonizes and integrates various geometric and semantic datasets from geological and geotechnical investigations to manage them in a data model based on GeoSciML and CityGML. A data exchange between CityGML and IFC is also possible, as shown in the study by Khan et al. (2023). The preceding examples illustrate the distinct focuses of IFC and CityGML data models. IFC is tailored to detailed building modeling, whereas CityGML is structured to support spatial data management across cities and neighborhoods. Additionally, the representation of the Level of Detail (LoD), geometric representation, and georeferencing of model objects in CityGML, as explained in Donaubauer et al. (2024), is better suited for this study than the implementation in the IFC format. Due to the focus of this study on urban planning, the GeoSciML and CityGML data model structures were combined as part of the data model design for the Stuttgart case study.

2.2 Urban Subsurface Modeling Workflows

In the context of 3D modeling, the model elements (geological boundary surfaces, volume objects for geological strata, buildings and infrastructure) are constructed from the base dataset collected by various investigations or transferred from previous models. Donaubauer et al. (2024) describe the different geometric representations of the various model elements that should be constructed in an integrated 3D subsurface model. The established workflow in building modeling involves the combination of various basic geometric shapes based on the Constructive Solid Geometry technique (CSG) with the objective of modeling complex volumetric elements, such as buildings and infrastructure elements. In the field of geological 3D modeling, it is common to construct model elements as boundary representations (B-Rep), where solid geometries are represented by their surrounding boundary surfaces. In a subsequent modeling step, it is also possible to extract solid geometries, which lie between the B-Rep modeling elements. The transformation of the 3D subsurface model into a voxel model requires a decomposition of B-Rep or solid geometries into uniform volumetric elements (Khan et al., 2023; Koch et al., 2017). Modeling software for geospatial analysis uses triangulation as an interpolation method to construct model elements as surfaces (e.g., geological strata as boundary surfaces) and volumes (e.g., geological strata as volumetric elements) from point-, line- and surface-based data (e.g., well logs, contour lines and profile sections). The interpolated model elements consist of a mesh of triangles for surfaces and tetrahedrons for volumes. The aim of constructing a voxel model is also to generate a grid or mesh by decomposing model elements. When modeling a voxel grid, the model geometries will be broken down into uniform volume cells known as voxels.

Case studies on subsurface modeling and the representation of subsurface information using voxel models have been presented in previous research. For example, Köbberich et al. (2022) demonstrated the modeling of bedding stiffness, a geotechnical parameter that describes the resistance of the subsurface to deformation within an IFC-compliant building model. The bedding stiffness calculated in a geotechnical model can be converted into the IFC format and subsequently assigned to the voxels of a building model. Khan et al. (2023) were able to construct a subsurface model for an urban area built from voxels that not only includes homogeneous strata but also represents inhomogeneous subsurface structures. The 3D subsurface model, including the semantic and geometric information managed in a data model, can be transformed between the IFC and CityGML formats.

2.3 Integrated Subsurface Data Management at the Geological State Surveys

When designing an integrated data model, the architecture of established information systems must also be taken into account. Geological Surveys, for example, have developed data models for managing incoming geological data. After the geological data are received, all relevant datasets should be retrievable for the modeling workflows of 3D models and other geodata analysis products based on the current state of the base data and geological modeling techniques. The final products are visualized through externally accessible geodata services and published as web services or downloadable data.

The State Authority for Geology, Mineral Resources and Mining – Department 9 of the Regierungspräsidium Freiburg (LGRB), which serves as the geological survey of the federal state of Baden-Württemberg, maintains the borehole and outcrop Database (ADB). This database provides fundamental data for geological modeling within the scope of the statewide geoscientific mapping program (Geologische Landesaufnahme, GeoLa), supporting disciplines such as geology, hydrogeology, engineering geology, economic geology, and soil science (Schmidt, 2015). As described in (Rupf and Nitsch, 2008), the LGRB has also constructed several regional and statewide geological 3D models based on this data. For external data provision, the LGRB operates a map viewer that is intended to enable the retrieval and display of the entire range of geodata services offered by the LGRB for external users. LGRB (2021) offers in the map viewer 39 different geodata services with over 550 topics from the various specialist departments of the LGRB.

Model elements from the 3D models of the state service are also available and can be analysed in detail with user tools such as cross-section tools or site assessment tools for the planning of shallow geothermal probes.

Systems for managing and visualizing geological 3D models in a 3D environment are also being developed at Geological State Surveys in Germany. GiGa provide a framework known as "Geoscience in Space and Time" (GST), developed for this purpose, which is already in use by the geological state services of Brandenburg, Bavaria, Saxony, Saxony-Anhalt, Hesse, North Rhine-Westphalia, and Lower Saxony (GiGa, 2025). Integrated subsurface models for geological and urban 3D modeling have also been implemented in GST. For example, Lehné et al. (2025) describe the implementation of a hydrogeological 3D model with integrated geometries of underground infrastructures and foundations for urban planning applications.

3. Modeling Workflow for the Unified 3D Model

3.1 Geological 3D Model Construction

The geological 3D modeling in this study is carried out with the geological software tool Aspen SKUA V 14.5 (SKUA). SKUA is a geological modeling software designed to facilitate the reconstruction, visualization, and analysis of geological structures and allows geoscientists and engineers to reconstruct subsurface geology, integrating diverse data types such as borehole information, geological maps and geoscientific surveys. Building on the concept introduced in Pusacker et al. (2024), this study continues work on the Stuttgart case study as highlighted in Fig. 1.

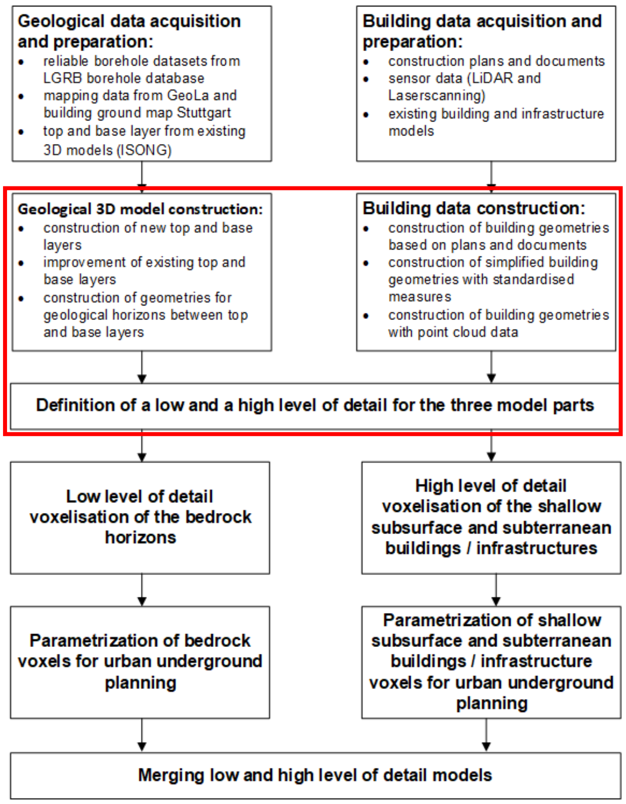

Figure 1. Diagram of the model workflow. The section marked in red represents the part of the modeling workflow that is investigated in this paper. Modified from Pusacker et al. (2024).

The study area (see Fig. 2), located in Stuttgart, the capital of the federal state of Baden-Württemberg, covers approximately 1,100,000 m², extending 960 meters in the north–south direction and 1,150 meters in the east–west direction. The planned 3D model will extend to a depth of 130 meters.

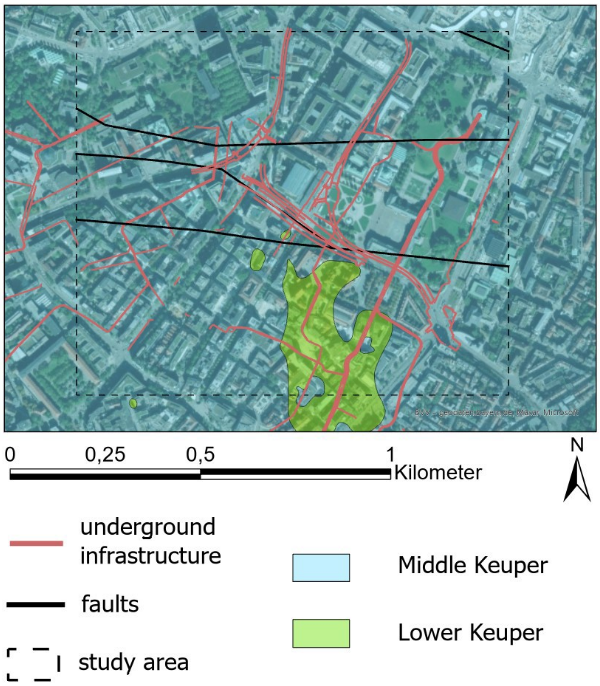

Figure 2. The study area in the city center of Stuttgart (black dashed line). The aerial photograph of the study area shows the densely built-up area within the study area, including residential buildings and public structures such as Schlossplatz, the "Neue Schloss", and the "Alte Schloss". The map also highlights the location of the underground infrastructure (red lines), the bedrock horizons located beneath the Quaternary topset beds (shaded in blue and green), and the course of the confirmed faults within the study area (black lines). (Geobasisdaten © Landesamt für Geoinformation und Landentwicklung Baden-Württemberg, www.lgl-bw.de, Az.: 2851.9- 1/19)

For the case study Stuttgart, the base layers of the bedrock strata Lower and Middle Keuper were modeled with the Discrete Smooth Interpolation method (DSI), which was initially introduced by Mallet (1992). The base layers derived from previously developed 3D models are designated as constraints to incorporate the results of earlier modeling efforts for the bedrock horizons. Further constraints are intended to incorporate the latest data on the depth of the base layers for interpolation. Therefore, layer boundaries derived from actual well log data in the ADB are added to the geological 3D model as point clouds.

Where the outcrop lines between the Middle and Lower Keuper are exposed at the surface, the boundary between these strata can be assumed. Accordingly, these line data, digitized from the Building Ground Map Stuttgart (BGM) (Rogowski, 2017), were used as constraints for the interpolation of the base of Middle Keuper.

The BGM also provides additional information on the depth of the Quaternary base layer. The thickness of the Quaternary topset bed is provided as a contour map, which is part of the BGM. With the contour lines and the well log data on the depth of the

Quaternary topset bed, the base layer of the Quaternary was re-interpolated. The distribution area of the unconsolidated rock layers subordinated to the Quaternary is also taken from the BGM. Their depth has been derived from the ADB well log data and is used to interpolate the depth level of the interpolated base layer of the unconsolidated rock layers.

3.2 Building Data Construction

At this stage of the study, the focus of the construction of building and infrastructure models is on the transformation of urban base data for integration into SKUA. Accordingly, this section will present how the urban modeling formats DWG and CityGML, as well as the sensor data format LAS, need to be prepared for import into SKUA (see Fig. 3). Various software tools are used for this data processing. In particular, the Feature Manipulation Engine (FME) offers a wide range of tools for converting urban and spatial data formats. Additionally, the GIS software tools QGIS 3.38 and ArcGIS Pro 3.2.2 were used for specific data conversions.

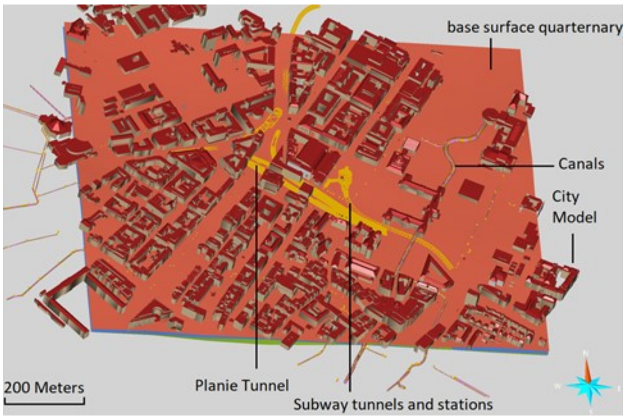

Figure 3. Location of infrastructure components, as described in section 3.2, within the Quaternary horizon (red) of the study area.

3.2.1 LAS-Workflow for the Integration of the Planie Tunnel: The workflow for processing the LAS dataset for the Planie Tunnel Stuttgart starts with thinning the point cloud datasets using the FME Point Cloud Simplifier tool to reduce data density while preserving essential details. The point spacing is typically in a centimeter range. For 33 LAS datasets, which encompass a section-wise divided LiDAR point cloud of the Planie Tunnel and the overlying terrain including buildings, vegetation, and above-ground infrastructure, the data volume could be reduced from 2.10 GB to 310 MB. The FME Point Cloud Coercer tool facilitates the conversion from a LAS dataset to a multipoint geometry. Moreover, FME can export this Multipoint Geometry as a shapefile. In this format, the point cloud datasets can be used in other software programs such as SKUA and ArcGIS Pro.

3.2.2 DWG Workflow for Sewer Systems and Subway Tunnels: The main sewer system (diameter > 2 m) and the subway tunnels, including the Börsenplatz and Schlossplatz stations, are each available as a DWG file. With FME, the DWG format can be converted into a shapefile with multipatch geometry. FME converts a single DWG file into multiple shapefiles for individual building components such as staircases, tunnel tubes, floors, etc., with sizes ranging from tens of centimeters to meters.

3.2.3 CityGML 3D Building Model Workflow: The city model for the study area Stuttgart is available as a CityGML file with LoD-2. LoD-2 describes all buildings with standardized roof forms, aligned according to the actual ridge line. The positional accuracy of the building footprints corresponds to that of the ALKIS cadastral data, which is within the centimeter range. The vertical accuracy of the city model is approximately 1 meter. For the import of the CityGML files into SKUA, a transformation of the data to the DXF format using FME was performed. Multiple DXF files, which consist of a large number of individual model elements, are created for the various components of the city model, such as buildings, walls, roofs, and terrain. The numerous individual model elements are inappropriate for further modeling workflows. Therefore, the DXF file created in FME is imported into ArcGIS. This software allows the export of more consolidated DXF files.

4. Development of the Integrated Data Model for the Case Study Stuttgart

For an integrated conflation of buildings, infrastructure, and geological structures in 3D voxel models for urban areas, existing data model standards need to be extended, and relationships between the features must be established (see Fig. 4). While well-known schemas for geological and urban data management, such as GeoSciML and CityGML, already support a variety of urban and geological data types, they require adaptation to address the complex requirements of a unified voxel 3D model. Such requirements are appropriate voxel sizes for different parts of the model depending on size, extent and spatial complexity of model elements as well as data availability for modeling. Additionally, the model should be built on an efficient data structure, such as an octree-based framework, to enable scalable generation, storage, and real-time management of large voxel datasets. These requirements are essential for maintaining performance and precision across diverse modeling tasks.

Within the framework of geological 3D modeling, the model elements (geological boundaries, volume objects for geological strata, buildings and infrastructure) were derived from the base data. Both the base data and the model features are to be managed within the data model introduced in this paper. Moreover, the information about the relationship between base data and modeling data should be preserved.

The core element of the conceptualized data model will be the constructed model objects. There are above-surface and below-surface model elements that can be captured and managed within the CityGML schema. Geological model elements can be described using GeoSciML data model structures. The data model is intended to gather the semantic data of the model elements as well as the information about their model geometries.

The semantic data model of geological features encompasses all relevant information associated with geological units, structural features and geomorphological formations represented by a model element. For the case study, the bedrock horizons, as well as the unconsolidated sediment layers, are taken into account as geological units. Verified faults in the study area are considered as structures in the case study. Well-known natural or anthropogenic geomorphologic features in the study area include sinkholes, landslides, the city moat, as well as building foundations and basements. The extension of the CityGML data model for the storage and exchange of geoscientific information is based on the study from Tegtmeier et al. (2014) extended with classes for voxel modeling.

https://doi.org/10.5194/isprs-annals-X-4-W7-2025-97-2025

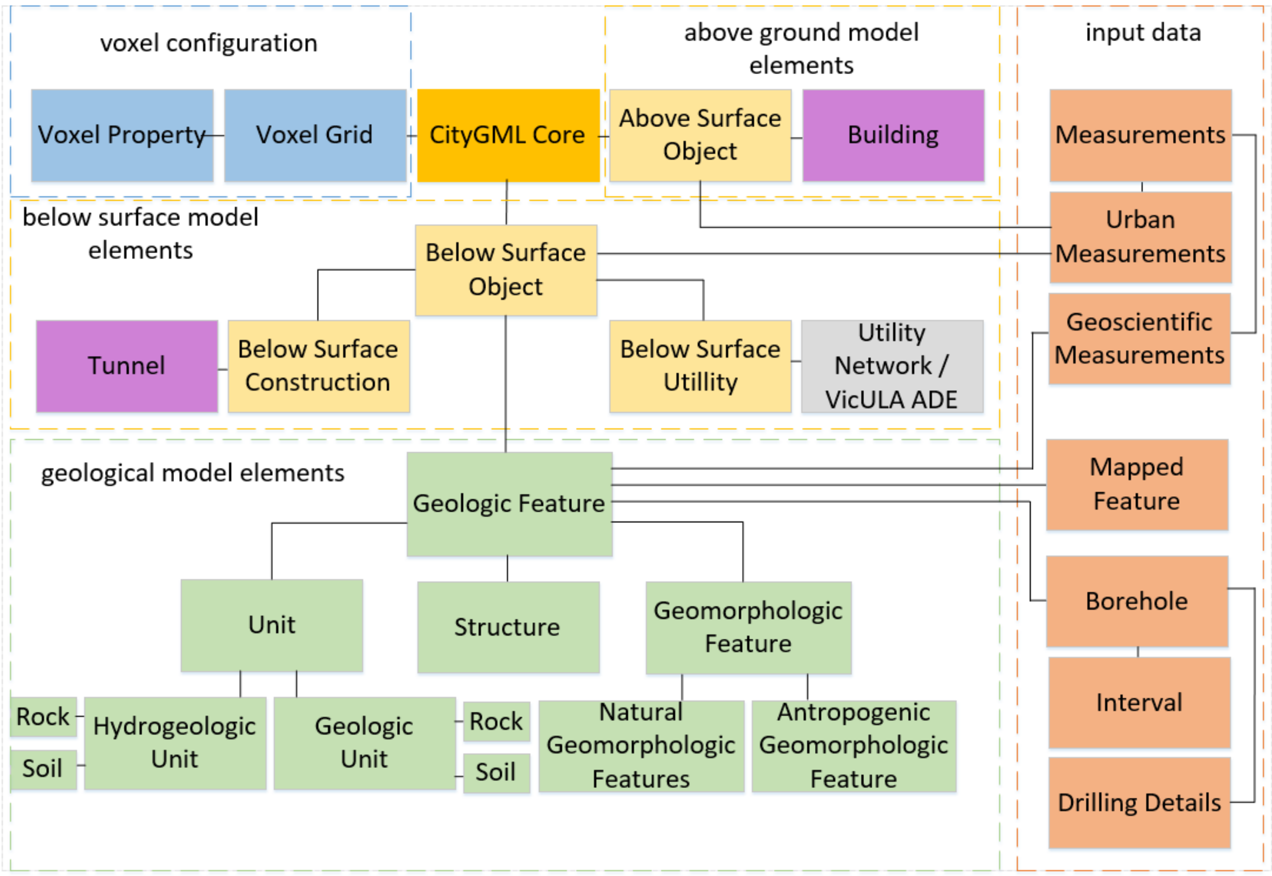

Figure 4. Class diagram of the 3D subsurface model based on application schemas for CityGML and GeoSciML. The modules of the CityGML core (orange) should be extended with specialized CityGML classes for above ground and underground modeling (yellow), which partly belong to the OGC-standardized modules (purple) or have been developed in the form of ADEs (grey). The GeoSciML classes include modules for managing geoscientific model elements (green) and the base data used for constructing these model elements (red). The configuration of the size, geographical extent and the resolution of the voxel grid is defined in the class Voxel Grid (blue). Parameters for modeling workflows should be managed in the subordinate class Voxel Properties.

Above-ground and below-surface model elements comprises data from city models and from the underground infrastructure. A section from the 3D building model for Baden-Württemberg in LoD-2 is added to the subsurface model as a representation for the buildings in the study area. The subsurface infrastructure is categorized into the classes Tunnel and Utility Network. Extension modules for the CityGML core are available for both the Buildings and Tunnels classes (Kolbe et al., 2021). Furthermore, the Utility Network and VicULA ADEs described by Bachert et al. (2024) and Dechamps (2024) are to be examined for their suitability as extensions for managing the data of canals and shafts. For input data, such as borehole data or geological maps, specifically tailored classes will be created that capture a broad range of information from the base data. The geological base data will be subordinate to the class Geologic Feature. For each Geologic Feature, it should be recorded which base data were used for modeling the model object. Urban model elements are often constructed directly, but they can also be generated based on sensor data. For example, buildings or infrastructure objects such as the Planie Tunnel in the study area are reconstructed from LiDAR data. These data can be managed within the Urban Measurements class.

The configuration of the voxel model is defined in a dedicated class within the data model. The associated Voxel Properties class assigns key modeling values to the used voxel grid. The study area introduced in Section 3.1 is discretized using a 3D voxel grid with a maximum resolution of 20 cm edge length per voxel. At this resolution, the voxel space covers 4,800 voxels along the X-axis, 5,750 voxels along the Y-axis, and 650 voxels along the Z-axis, resulting in a total of approximately 17.94 billion voxels. This 3D voxel space is divided into cells of 64 voxels per axis and, for the study area, and generates a grid of 67,500 cells. Each cell is stored as a record in a column of a relational table. A cell is not represented as a regular grid of 64 voxels per axis, with 262,144 separate values, but rather as an octree data structure, a hierarchical data structure of nodes at seven levels (0 to 6). A node at level L corresponds to a cubic collection of voxels (see Table 1).

Level (L)	Subdivisions per Axis (2^L)	Voxels per Node (2^{3L})	Voxel Edge Length
0	1	1	0.2 m
1	2	8	0.4 m
2	4	64	0.8 m
3	8	512	1.6 m
4	16	4,096	3.2 m
5	32	32,768	6.4 m
6	64	262,144	12.8 m

Table 1. Corresponding number of voxels, voxel size and cube size per node, assuming a voxel edge length of 0.2 m at level 0.

The database uses a variable-length field to store octree data, allowing efficient compression in homogeneous areas by storing only one high-level node per cell. Aggregation strategies differ by data type: volumetric objects use majority voting, while physical measurements like temperature or soil moisture use the median. Specialized functions are applied to preserve the integrity of surface, line, and point features at coarser resolutions.

5. Assessment of the Geological 3D Model for Use in Urban Subsurface Modeling

The construction of the geological 3D model aims to derive model elements from the geological base data that both represent the general geological framework of the study area and incorporate complex structures that reflect the specific geological conditions in the city of Stuttgart. An enhanced reconstruction of the boundary layers between the solid rock horizons of the Lower and Middle Keuper was accomplished by incorporating the tectonic structure of the study area through the subdivision of the bedrock into individual fault blocks. The originally horizontal bedrock horizons have been displaced relative to one another due to deformation processes along subsurface fractures. By modeling the layer boundaries for each fault block, displacements along the faults are captured more accurately. The assumption of a fault-block system, as illustrated in Fig. 5, is consistent with the geological interpretation of tectonics in the Stuttgart region (Geyer et al., 2023). Stuttgart is located to the east of the Filder Graben. During the formation of such a graben structure, the subsurface in the graben area subsides and is affected by extension. As described by Ufrechter (2018), this dilation leads to the development of a fault system along the margins of the Filder Graben. It should be emphasized that the segmentation of the study area into fault blocks is a simplification of the actual tectonic structure of the subsurface. The fault-block tectonics are significantly overprinted by subsequent tectonic activity and by subsidence resulting from gypsum leaching (Ufrechter, 2018; Rogowski, 2017). As emphasized by Rogowski (2017), the subsidence within the Middle Keuper limits the reliability of data from this unit. Consequently, only the positions of the Lower Keuper or the underlying Muschelkalk can be considered as reference horizons in the context of the BGM data. However, the available data on the displacement of geological strata at this depth is limited. Due to this, only the faults verified in the Geola datasets were used to reconstruct the fault block system.

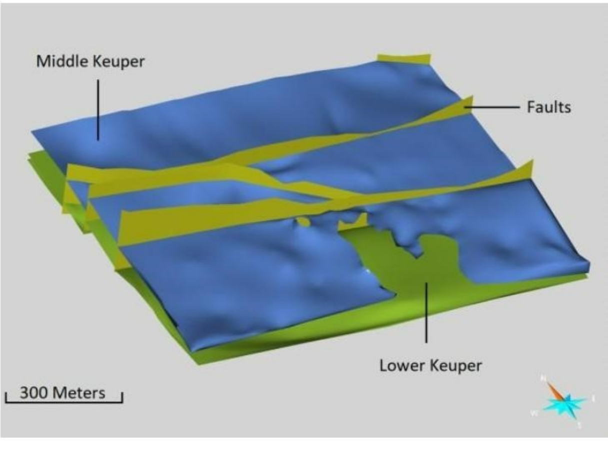

Figure 5. The base of the Middle Keuper (blue) and the Lower Keuper (green) are represented as boundary surfaces of the bedrock horizons. The faults of the fault-block system have been modeled as vertical surfaces (yellow).

The Quaternary Floodplain Sequence of the Nesenbach and Vogelsangbach in the study area (see Fig. 6) consists of various unconsolidated rocks that were deposited in different depositional environments and have varying building ground properties. Floodplain Gravel, Cut-off Lake Sediments, and Floodplain Clay are fluviatile sediments transported by the streams in the study area. The Floodplain Gravel consists of coarser sediments with a grain size from 2 mm to 200 mm. These gravel and block sediments eroded from the Keuper, Muschelkalk, and Lower Jurassic bedrocks were deposited in the valley channels of the Nesenbach and Vogelsangbach valleys. The Cut-off Lake Sediments comprise organic deposits which formed in stagnant water bodies. In the oxygen-poor environment of these water bodies, Cut-off Lake Sediments developed from undecomposed organic material, partially interbedded with fine sand and clay. Floodplain Clay is a fine-grained sediment accumulated in floodplains, where it could accumulate slowly in low-flow areas of the watercourses. This sequence of sediments could be reconstructed for the shallow subsurface based on the available input data.

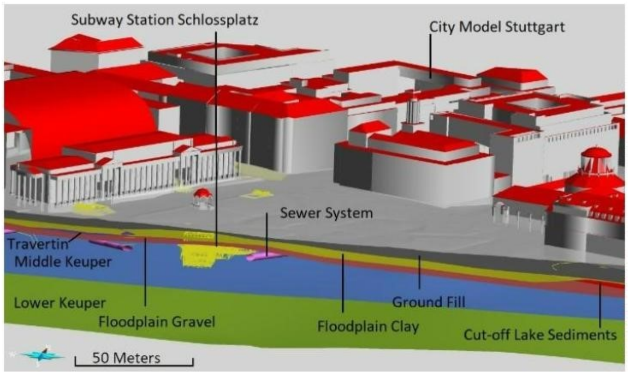

Figure 6. Unified 3D model of the subsurface beneath Schlossplatz. The positions of the bedrock horizons are represented by the base surfaces of the Middle Keuper (light blue) and Lower Keuper (green). The unconsolidated rock layers Ground Fill (grey), Floodplain Gravel (pink), Travertine (dark blue), Cut-off Lake Sediments (red), and Floodplain Clay (dark yellow) represent the subdivision of the topset beds. The Schlossplatz subway station (yellow grid) and the sewer system (purple pipes) depict the underground infrastructure in the model section. The buildings above the terrain surface are taken from the Stuttgart city model

The genesis of Travertine is not related to fluvial sedimentation. The calcareous sands and silts of the Travertine are formed as a result of the proven ascent of mineral waters in the study area. The carbonated mineral waters dissolve calcium carbonate from the calcareous bedrock horizons and transport it to the surface. At the surface, the degassing of CO_2 from the mineral water causes the precipitation of calcium carbonate. The naturally deposited Quaternary unconsolidated rock layers are overlain by anthropogenically accumulated Ground Fill material.

6. Assessment of Octree Voxel Model

The reduction in storage requirements and the performance improvements achieved through the octree data structure introduced in Section 4 can be demonstrated by comparing it to conventional storage methods. An octree data structure model allows users to store data at coarser levels (e.g., level 2 with 0.8 m resolution instead of the finest 0.2 m resolution) to reduce data volume, while still enabling queries at finer resolutions by referencing values from higher-level parent nodes. When new data is added, the system automatically aggregates values upwards and simplifies the structure by removing lower-level nodes that duplicate their parent's value. This recursive compression ensures that only exceptions to inherited values are stored, which greatly improving storage efficiency. Besides providing significant storage reduction, octrees bring better processing speed of certain operations. A Overlay of voxels, for example, works the same way on input octree nodes of any size. The ability to query at higher (coarser) levels is significant for visualisations ("zoom out") and for submitting voxel datasets to external software that may be limited by dataset size. In summary, all voxels in the entire cell will always "exist" at all levels, but the number of nodes will differ from cell to cell. In the Stuttgart case study, the total storage space was calculated for the seven layers: canals, LiDAR point cloud, Buildings, Quaternary Horizon, Middle Keuper Horizon, Lower Keuper Horizon and Deeper bedrock, by summing the size of all octree units, resulting in 1.2 GB. As shown in Table 2, volumetric layers could be significantly compressed, with storage requirements dropping to less than one byte per voxel on average. However, no reduction was possible for the LiDAR point cloud due to the highly irregular and sparse distribution of its points.

Layer	Octree Size (Bytes)	Voxels	Bytes per Voxel
Canals	5,441,592	14,583,649	0.37313
LiDAR	33,085,424	7,792,423	4.24597
Building	233,438,040	1,028,927,454	0.22722
Quarternary	125,604,564	790,965,845	0.16811
Middle Keuper	206,312,288	1,997,053,371	0.10429
Lower Keuper	209,710,360	2,559,419,247	0.08194
Deeper	363,562,616	4,315,187,740	0.02232
Total	**1,177,154,884**	**10,713,929,729**	**0.10987**
Combined	**580,638,568**	**10,636,236,660**	**0.05459**

Table 2. The table presents voxel statistics for seven individual layers and total statistics for the complete 3D model. The upper seven rows show data for separate layers. The row "Total" summarize the results for the single layers. The row "Combined" merges all layers into one, assigning each voxel a unique value, eliminating overlaps.

In contrast to this, representing the Stuttgart study area as a full 3D array at 20 cm resolution would require storing all 17.94 billion voxels, including those with zero values, separately for each of the seven layers. This results in a total of 125.6 GB with each voxel occupying one byte of storage. In integer point clouds, only non-zero-value voxels are stored. In addition to the assigned voxel value, the x, y, and z coordinates are also needed, increasing the storage to four bytes per voxel. Thus, for the study area, 42,544,946,640 bytes (≈ 42.5 GB) of storage space would be required.

7. Conclusion

This paper presents a data model schema tailored to urban and geological subsurface models based on voxel representations. The schema is designed to manage both semantic data and model elements in an integrated manner. The review of the current state of the CityGML Core and the existing ADEs for subsurface infrastructure management has confirmed that urban and sensor data can be managed within CityGML data model structures. With the additional use of GeoSciML, also geoscientific datasets can be managed. The central element of the data model is the set of model elements constructed based on interpolation of the input data. As shown in the modeling workflow described in Section 3.1, the base data serve as essential constraints required for the reconstruction of the model. For this reason, the data model was designed to ensure traceability of the data used in the construction of the geological model elements. The geological modeling software SKUA was successfully tested as a modeling environment for integrating urban and geological data. SKUA enables the modeling of model elements composed of voxels. Additionally, SKUA offers tools to assign parameters, such as material properties, to voxels. These parameterized voxels can be used in SKUA, for instance, for geoscientific process modeling (Aspentech, 2022). It was shown that varying voxel resolutions, from decimeter to meter scale, are required to accurately capture the high LoD of building elements, while simultaneously optimizing storage by using coarser resolutions for geological horizons with lower LoD. For advanced model analyses, such as assessing the accuracy of datasets, even higher resolutions may be required. This approach allows a balanced trade-off between accurate modeling and data efficiency depending on the model component.

Acknowledgement

We sincerely thank the Civil Engineering Department of Stuttgart for their support of this research. In particular, we extend our gratitude to Mr. Silvester Koci and Mrs. Silvia Deist, whose provision of building and infrastructure data for the city center' was essential to the development of this concept.

References

Aspentech, 2022. Aspen SKUA™, https://www.aspentech.com/en/resources/brochure/aspen-skua (20 April 2025)

Bachert, C., León-Sánchez, C., Kutzner, T., Agugiaro, G. 2024. Mapping the CityGML Energy ADE to CityGML 3.0 Using a Model-Driven Approach. *ISPRS International Journal of Geo-Information, 13(4), 121.* https://doi.org/10.3390/ijgi 13040121

Donaubauer, A., Kaden, R., Seuß, R., Grubert, U., Kolbe, T. 2024. CAD, BIM und GIS – digitale Modelle der gebauten Umwelt. *Leitfaden Geodäsie und BIM. Version 4.0, Bühl/München, 69-85.*

Building Smart, 2022. IFC 4.3.2.0 – Introduction, https://ifc43-docs.standards.buildingsmart.org/IFC/RELEASE/IFC4x3/HTML/content/introduction.htm, (25 April 2025)

Chapman, D., Providakis, S., Rogers, C., 2020. BIM for the Underground – An enabler of trenchless construction. *Underground Space, 5, 354–361.* https://doi.org/10.1016/j.undsp.2019.08.001

Commission for the Management and Application of Geoscience Information (CGI), 2021. GeoSciML v4.1, http://www.geosciml.org/doc/geosciml/4.1/documentation/html (18 April 2025).

Coors, V., Pietruschka, D., Zeitler, B., 2022. iCity. Transformative Research for the Livable, Intelligent, and Sustainable City: Research Findings of University of Applied Sciences Stuttgart. *Springer Nature, 282-271.* https://doi.org/10.1007/978-3-030-92096-8

Dechamps, L., 2024. Spatial and semantic enrichment of utility networks data, Doctoral diss., Delft University of Technology.

Geyer, M., Nitsch, E., Simon, T., 2003. *Geologie von Baden-Württemberg. E. Schweizerbart, 254- 295*

Giga Infosystems, 2025., GST Informations, https://giga-infosystems.com/references (25 April 2025).

Jeong, D., Lee, C., Choi, Y., Jeong, T., 2024. Building digital twin data model based on public data. *Buildings, 14, 2911*. https://doi.org/10.3390/buildings14092911.

Kasprzyk, J. P., Nys, G. A., Billen, R., 2024. Towards a multi-database CityGML environment adapted to big geodata issues of urban digital twins. *Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci. 48, 101-106.* https://doi.org/10.5194/isprs-archives-XLVIII-4-W10-2024-101-2024.

Khan, M. S., Kim, I. S., & Seo, J., 2023. A boundary and voxel-based 3D geological data management system leveraging BIM and GIS. *Int. J. Appl. Earth Obs. Geoinf.,* 118, 103277. https://doi.org/10.1016/j.jag.2023.103277.

Koch, C., Vonthron, A., König, M., 2017. A tunnel information modelling framework to support management, simulations and visualisations in mechanised tunnelling projects. *Autom. Constr. 83, 78–90*. https://doi.org/10.1016/j.autcon.2017.07.006.

Komar, T., James, P., 2024. Orchestrating Urban Footfall Prediction: Leveraging AI and batch-oriented workflow for Smart City Application. *Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci. 48, 119-124.* https://doi.org/10.5194/isprs-archives-XLVIII-4-W10-2024-119-2024.

Kolbe, T. H., Kutzner, T., Smyth, C. S., Nagel, C., Roensdorf, C., Heazel, C., 2021. OGC City Geography Markup Language (Citygml) Part 1: Conceptual model standard standard approved.

Köbberich M., Schneider O., Schildknecht L., Gafner T., Volken S., Wehrens P., Preisig M., Schwab P., Staub B., Grischott R., Figi D., Wirth S., Löpfe R., Roth S., Thüring M., Stockermeyer M., Graf J., Huber M., Möri A., 2022. Innovationsprojekt GEOL_BIM – Abschlussbericht. Schweizer Geologen Verband.

Landesamt für Geologie, Rohstoffe und Bergbau (LGRB), 2021. LGRB-Kartenviewer, https://maps.lgrb-bw.de/ (April 14 2025)

Lehné, R., 2025. GST at HLNUG, GST User Meeting 2025, https://giga-infosystems.com/wp-content/uploads/2025/03/GST_at_HLNUG_2025.pdf (25 April 2025)

Mallet, J. L., 1992. Discrete smooth interpolation in geometric modelling. *Computer-aided design, 24(4), 178-191.* https://doi.org/10.1016/0010-4485(92)90054-E.

Niu, T., Lin, B., Zhou, L., Lv, G., 2024. A 3D bedrock modeling method based on information mining of 2D geological map. *Earth Sci. Inf., 17.5, 4067-4094.* https://doi.org/10.1007/s12145-024-01375-7

Otón, M. P., González, L. R. C., Van Campenhout, I., Schokker, J., Watson, C. & van der Meulen, M. J., 2021. Emerging user needs in urban planning. - *Applied Multidimensional Geological Modeling: Informing sustainable human interactions with the shallow subsurface, 383-401.* https://doi.org/10.1002/9781119163091.ch16.

Padsala, R., Gebetsroither-Geringer, E., Peters-Anders, J., Coors, V., 2021. Inception of Harmonising Data Silos and Urban Simulation Tools Using 3D City Models for Sustainable Management of the Urban Food Water and Energy Resources. *ISPRS Ann. of Photogramm. Remote Sens. Spatial Inf. Sci., 8, 81-88.* https://doi.org/10.5194/isprs-annals-VIII-4-W1-2021-81-2021

Pusacker, K., Coors, V., Eckhardt, J. D., Rupf, I., 2024. A Concept for 3D Geological and Urban Subsurface Modeling with a Unified Voxel Model Examined by a Case Study for the City Center of Stuttgart (Baden-Württemberg), Germany. *ISPRS Ann. Photogramm., Remote Sens. Spatial Inf. Sci., 10, 193-200.* https://doi.org/10.5194/isprs-annals-X-4-W4-2024-193-2024

Qu, Y., Perrin, M., Torabi, A., Abel, M., Giese, M., 2024. GeoFault: A well-founded fault ontology for interoperability in geological modeling. *Computers & Geosciences, 182, 105478.* https://doi.org/10.1016/j.cageo.2023.105478.

Rogowski, E., 2017. Der Baugrund von Stuttgart - Erläuterungstext zur digitalen Baugrundkarte. Landesamt für Geologie, Rohstoffe und Bergbau Baden-Württemberg.

Ran, X., Xue, L., Sang, X., Pei, Y., Zhang, Y., 2022. Intelligent generation of cross sections using a conditional generative Adversarial Network and application to regional 3D Geological Modeling. *Mathematics, 10, 4677.* https://doi.org/10.3390/math10244677

Rupf, I., Nitsch, E., 2008. Das Geologische Landesmodell von Baden-Württemberg: Datengrundlagen, technische Umsetzung und erste geologische Ergebnisse. *LGRB-Informationen, 21, 81.*

Schmidt, C., 2015. LGRB goes 3D. *GeoLa-Meeting 23.04.2015, Freiburg im Breisgau.*

Tegtmeier, W., Zlatanova, S., Van Oosterom, P. J. M., & Hack, H. R. G., 2014. 3D-GEM: Geo-technical extension towards an integrated 3D information model for infrastructural development. *Comput. Geosci. 64, 126-135.* https://doi.org/10.1016/j.cageo.2013.11.00.

Ufrechter, W., 2018. Die Tiefengrundwässer im Oberen Muschelkalk zwischen Albvorland und oberschwäbischem Molassebecken—ein Hydrogeologisches Modell. *Schriftenreihe des Amtes für Umweltschutz, Stuttgart, 1, 7-158.*

Virtanen, J. P., Alander, J., Ponto, H., Santala, V., Martijnse-Hartikka, R., Andra, A., Sillander, T., 2024. Contemporary development directions for urban digital twins. *Int. Archives Photogramm. Remote Sens. And Spatial Inf. Sci., 48, 177-182.* https://doi.org/10.5194/isprs-archives-XLVIII-4-W10-2024-177-2024

Wu, J., Chen, J., Chen, G., Wu, Z., Zhong, Y., Chen, B., Huang, J., 2021. Development of data integration and sharing for geotechnical engineering information modeling based on IFC. Advances in *Civil Engineering, 2021(1), 8884864*. https://doi.org/10.1155/2021/8884864

Zhang, Z., Wang, G., Carranza, E. J. M., Liu, C., Li, J., Fu, C., Liu, X., Chen, C., Fan, J., Dong, Y., 2023. An integrated machine learning framework with uncertainty quantification for three-dimensional lithological modeling from multi-source geophysical data and drilling data. *Engineering Geology, 32* https://doi.org/10.1016/j.enggeo.2023.107255