Towards a collaborative and interoperable 3D Building database – A case study in Walloon region

R. Hajji¹ and R. Billen²

¹ HassanII Institute of Agronomy and Veterinary Sciences, BP 6202-Instituts, 10101 Rabat, Morocco
² University of Liège, Unit of Geomatics, 17 allée du 6 Août, B4000 Liège, Belgium

Abstract. The definition of a 3D Reference model is the prerequisite for increasing data interoperability and facing regional, national and international challenges around geographic information. Such action requires investigation of 3D user requirements and a collaborative framework to reach a consensus on common 3D data specifications. The paper presents premise reflexions about relevant issues to shape efforts towards a methodological and generic approach for dealing with a collaborative 3D reference model as a fundamental building block of 3D GIS collaborative solution. As a pilot project, we demonstrate, through a case study of the Liege city in Belgium, how data collected from different providers in Walloon region can be reengineered and then integrated in a 3D collaborative interoperable database compatible with CityGML.

1. INTRODUCTION

For long decades, government agencies used to monopolize production and distribution of geographic information. A similar range of products such as cadastral and topographic maps was developed in many countries. Maps were not necessarily a consumer product [1]. Nowadays, the availability of Information and Communication Technologies (ICT) at affordable prices as well as the emergence of social web platforms like OpenStreetMap have demonstrated the potential to revolutionize the way geospatial data can be acquired and shared. Given these facts, a question can be asked: Will government agencies be attributed a new role: as a regulator actor instead of principal producer?

Currently, there is an increasing need for representing and analyzing the 3D world. Virtual globes are certainly optimized for exploration and browsing purposes [2], but cannot deal with a large spectrum of use cases. Only a 3D well-structured model can support “intelligent” spatial analysis and processes. So, potential producers are likely to balance economic interests, with considerations of increasing the availability and accessibility of 3D geographic information through the development of 3D GIS. The focus of interest about a 3D GIS is about 3D spatial analysis which are very relevant in many applications like emergency responses, geology, 3D cadastre and urbanism. It is now time to upgrade models, data acquisition methods, formats and products to meet new requirements. As it has been stated by Goodchild, the issues of a 3D GIS are more than technological [3]. Policies and national agreements through a collaborative framework are needed to deal with access and use of heterogeneous and distributed data sources.

Knowing that each audience has a distinct set of 3D data needs, a major concern is how to establish one integrated solution to reconcile divergent and not yet well identified requirements for 3D information. We believe that defining and implementing a complete 3D reference model would be challenging, money and time consuming. A step-wise process may be more appropriate to reach basic requirements which can be enriched and completed since 3D reference geographic information...
is available. Furthermore, a participatory approach is a powerful solution to face the challenge through establishing a 3D collaborative model.

In the following, we first analyze the main concepts and issues related to adopting a 3D reference model, and then we present a conceptual view of our collaborative vision. Finally, we explain our proposed methodology for evolving from 2D multi source data maintained in actual databases to reconstruct 3D information and populate a 3D collaborative database compatible with CityGML. Our fundamental objective is to demonstrate, through a case study, how a collaborative and interoperable 3D database can be established and populated.

2. TOWARDS 3D COLLABORATIVE REFERENCE MODEL: A CONCEPTUAL VIEW

2.1 “Reference” issue

First, the notion of “reference” associated to data has to be clarified in view of other concepts like ‘core’ and ‘fundamental’ data. Indeed, according to the SDI cookbook authors, the ‘reference’ data represents data that can be used to relate or ‘refer’ external information to the real world such as infrastructure theme, terrain elevation, hydrographic or abstract features like administrative boundaries, cadastral parcels and postal addresses [1]. Here, the reference data plays for a geographic information user the same role as a geodetic frame does for cartographers and surveyors. On the other hand, ‘core’ data constitutes a common denominator of all geographic information data sets by representing fundamental data that can be shared by most applications, with specifications compatible to those of reference data [1]. In this paper, we consider the concept of ‘reference’ data meets the last definition but we prefer using the ‘reference’ term to refer to fundamental 3D data that potential producers must make available to deal with a large range of 3D applications. To reach this objective, it seems logical to investigate and rely on users requirements [4], to define the types of real world objects the 3D reference model must represent and to choose an appropriate design of an integrated model capable of maintaining all the components of the geometric representation of real world objects in the same database [5]. Finally, we argue that ‘reference’ data concept is relative and associated to fitness of use and then dynamic if we consider that user requirements become more sophisticated when achieving a high degree of maturity.

2.1.1 3D user’s requirements

The first raised question in developing 3D GIS is about user’s requirements. In reality, 3D requirements are not well known by users themselves. Indeed, confronted by organizational, financial and market problems, users can hardly identify revolutionary new requirements [4]. Although carrying experiments and ideas [4], they believe 2D technologies sufficient to complete their tasks and then are unable to arise from relevant needs. Zlatanova stated that user’s 3D requirements can be investigated through a technology or an application-driven approach. The former studies the question regarding improvement made in new technologies dealing with 3D information while the later relies on user’s needs. A mix of the two approaches is more appropriated. Users can better understand their needs if they are aware of available 3D geo-information and the supporting technologies [6].

2.1.2 3D Reference objects

3D real objects of interest for urban area are application depending and then different types of objects may be more (or less) relevant than others [6]. Many authors have already addressed the question by identifying main relevant 3D objects for urban applications. As they were been reported by Zlatanova, studies have resulted to convergent classes of interest: buildings, terrain, vegetation, traffic network, public utilities and telecommunications are the main objects to be maintained by a 3D GIS. Underground infrastructures are also important to handle underground urbanism applications [4, 7].
The same author has distinguished four basic groups of real objects: juridical objects (e.g. individuals, institutions, and companies), topographic objects (e.g. buildings, streets, and utilities), fictional objects (e.g. administrative boundaries) and abstract objects (e.g. taxes, deeds, incomes) [4, 7]. Unlike the 3D topographic objects which are basically the 3D spatial objects currently maintained in a variety of information systems, non-spatial objects like abstract or fictional objects are still in the conceptual level. Even if the third dimension represents the reality, the needs of 3D information have been born and grown because of the limits of 2D data to deal with some application requirements. So, it seems obvious that basic objects of a 3D model are constructed by extending those of traditional 2D models (from 2D to 3D thinking) such as buildings, vegetation, city furniture and so forth.

2.1.3 Multi scale representation

Considering that several 3D applications may share a unique 3D reference database, dealing with multi representations is a fundamental issue. Some of the questions to make up a representation would be: what information to keep? How it is described? How it is organized in terms of data-structures? How it is coded? [8]. Depending on application requirements, a certain spatial and thematic granularity is needed. As it is defined in the CityGML specifications document, the concept of level of detail allows representing objects with regard to different degrees of resolution. It is characterized by differing accuracies (described as standard deviation of the absolute 3D point coordinates) and minimal dimensions of features [9]. Objects become more detailed with increasing LoD regarding both geometry and thematic differentiation [10].

According to the CityGML specifications associated to each level of detail, without regards to the quantitative aspect, the accuracy and the description of how rich would be a representation of a building in each level of detail might be confusing. Such as an example, a building with structured internal details as it’s required for LoD4 might be classified in a lower LoD because a lack of accuracy, which can cause a loss of information. In our opinion, the definition of the level of detail should be more sophisticated. The user should have the choice to define which type of level of detail (‘semantic’ or a ‘geometric’) is the most appropriated for the application. We note that a ‘semantic’ LoD is related to a semantic resolution which represents the precision of description. Certainly, for many applications like emergency services and urban development, a high semantic LoD is more privileged than a geometric LoD, but it’s recognized that acquiring data for representing an urban area with a high LoD is expensive, complex and time consuming. As an example, handling the complex internal structures of buildings is stymied by a lack of a cost-effective technology for indoor positioning that is comparable to GPS [3].

Finally, as mentioned in the introduction, potential producers may adopt a step wise process to establish 3D urban models. We think that having a large area with simplified and updated 3D model is a good realization to deal with basic needs and to offer a first 3D product which can be more extended and enriched by users. We believe that a complete global model to fill all needs does not exist. Thus, spending too much time in a planning stage to reach a coherent and complete model before implementing might not be an appropriate solution.

2.2 “Collaborative” issue

To respond to emergent challenges and responsibilities related to geospatial data at regional, national and also international scale (in particular Inspire directives), many actors must collaborate effectively to build National Spatial Data Infrastructure (NSDI). Technical, standard, and policy deficiencies result in time and effort losses on data production, management, and sharing [11]. In order to avoid duplicated efforts and expense occurring throughout jurisdictions at all levels of administration [12], data provided by many contributors must be integrated to develop reference data sets [1]. Doing so, data interoperability can be greatly improved [1]. To reach the objective of establishing a national 3D data reference, a geographic data framework must be established through a national agreement on content and
Usage, Usability, and Utility of 3D City Models

specifications. The framework will constitute a collaborative ‘datum’ on which organizations can build by adding their own detail and compiling other data sets [1]. As it was explained in our paper (submitted for Sageo, 2012) [13], upgrading to a 3D collaborative model would be a promising solution to anticipate interoperability and data consistency problems and also to establish common specifications for 3D data co-production. However, integrating existing data with geometric and semantic divergences is the first challenge to establish the basis of the 3D collaborative model. The degree of difficulty depends on how rich are the source schemas with regards to the target one adopted for the 3D collaborative database.

Recently, data crowd sourcing in forms of Voluntary Geographic Information (VGI) is an emergent source of geographic data that presents interesting research challenges [3, 14]. However, many authors have raised the issue about its credibility, quality and longevity as well as the mechanisms to make it benefic [3, 14–16]. Crowd sourcing might be investigated with regards to established specifications of a 3D collaborative model. Until no framework is available to study the adoption and use of VGI, usage can be restricted to generating updating alerts. Also, a benefic solution can be adopted by collecting VGI from a network of professionals. In all cases, potential data producers should motivate public by allowing free access in products and services.

In Belgium, many divergent initiatives have been conducted for each political region to establish Topographic Inventories (TI) with divergent models and different technical characteristics (resulting to PICC in Walloon region, URBIS in Brussels capital region and GRB in Flanders). Also, the National Geographic Institute (NGI) maintains and distributes a large set of geographic data and cartographic products. Actually, many individual initiatives are launched to upgrade the existing 2D Topographic Inventories into 3D Geospatial databases. Some cooperation’s aspects exist but are not formalized enough to be really efficient. There is a strong need to share resources in order to define a 3D collaborative model. Some of the raised questions are: How the existing data content may be enhanced and adjusted to match a national or global framework specification? How good are the 3D standards like CityGML for 3D modeling? And how both 2D and 3D databases can cohabit? A lack of a conceptual and methodological approach to deal with these fundamental issues constitutes a major obstacle.

We do recognize the high ambition to address the issue but our work aims at capitalizing on major accomplishments of research to propose a methodological and generic approach for dealing with a collaborative 3D GIS solution and to demonstrate, through a case study that we present in the next section, how data collected from different providers (in Walloon region) can be integrated in a 3D collaborative interoperable database.

3. CASE STUDY: TOWARDS A 3D COLLABORATIVE BUILDING DATABASE OF LIEGE CITY-BELGIUM

3.1 Basic objectives

The objective of the experiment in progress of a PhD research aims at investigating how existing multi source 2D data with Z information can be reused and integrated in a 3D collaborative database compatible with CityGML. This experiment must be seen as a pilot project of a collaborative approach towards a 3D GIS solution. Through a case study of the city of Liege, we aim at learning for future recommendation for a 3D GIS collaborative approach and also, to test an umbrella of 3D technologies to validate conceptual solutions. We note that our research is limited to building features. In this paper, only the geometric process is presented. The semantic integration is not addressed.

3.2 Used data

Two different data sets are used in this experiment. The first one has been provided by the SPW (Service Public de Wallonie) as map sheets of the PICC (Projet Informatique de Cartographie Continue) in a 3D
Figure 1. Processing Workflow.

shape file format where each X, Y point has a value of Z. The level of the Z measurement is indicated in the specification’s document. The second data set has been provided by the NGI in an Esri geodatabase file format. In the two data sets, buildings have been restituted by photogrammetry at the cornice level. This data was then imported and reprojected in the same coordinate system (Lambert Belge 1972) in the Arc GIS 10 environment.

3.3 Processing workflow

The workflow is illustrated in (Figure 1) which main steps are described in the following paragraphs.

3.3.1 Building heights interpolation

Before any process on multi source data takes place, we had to interpolate the building’s heights for each data set. To accomplish this task, a Digital Surface Model (DSM) and a Digital Terrain Model (DTM) were established using the ArcGIS 3D Analyst TIN Creation tool. The DTM was created from cloud points, axis layers and a selection of terrain characteristic lines by defining how features should participate at the triangulation process (mass points or hard lines). The DSM or exactly a ‘Digital Cornice Model’ (DCM) was generated from the contours of building’s roofs. After rasterization of the two TINs surfaces and the building contours roofs at a resolution of 2 m (chosen according the dimensions of the smallest building to be represented), statistical computing were used to extract the height of each building (Figure 2). “Zonal Statistics” function (of ArcGIS) was applied, first to the DTM and buildings to extract the minimum as the base height and next to the DSM and buildings to extract the maximum as the roof height. The difference represents the interpolated height of each building which a unique attributed identifier was generated in ArcMap. We note that the quality of the results can be evaluated using an external control like surveying methods to have heights of a representative sample of buildings.

3.3.2 3D Building Model

As we have stated in part (2) of this paper, collaboration of a large number of stakeholders, providers and users is fundamental to reach a national agreement on a standardized 3D Geo-information model. So, it is obvious that defining a 3D reference model is beyond the scope of this research. Our contribution is rather technical.

The international 3D standard: CityGML is a common information model for the representation of 3D urban objects. It plays a leading role in the modularization of urban geospatial information [9, 17]. But, its complexity makes it is hard to implement all its specifications [17]. For our experiment,
Usage, Usability, and Utility of 3D City Models

Figure 2. Building heights interpolation.

Figure 3. 3D Building Model.

we have adopted a simplified version of CityGML based on the building thematic module (Figure 3) where buildings are represented as blocks in LoD1 with additional information about boundary surfaces (Roof/Wall/Ground Surface). We can say that buildings are represented in LoD2 with flat roofs.

3.3.3 Data integration

Data integration is recognized as a very complex process that researchers still accord much attention. It is conducted through three main steps: 1) The pre integration which consists of a good understanding of the content of each database and the rearrangement and mapping between models to show similarities and possible connections; 2) The correspondences investigation: through an identification and declaration of correspondences between the elements of the schemas and the geometrical instances of the databases and finally 3) The integration by defining explicit rules to translate and restructure the initial schemas and data transfer to the new system (See [18] for a detailed description). In our experiment, we have to deal with multi source data, so a data matching process is used to select the appropriate one is required.
3.3.3.1 Correspondences investigation: Data matching

Since the process presented here is geometry-based, a semantic schema matching between the source model and that of CityGML is not addressed here. The objective is focused on a reconstruction of 3D geometric building features according to the CityGML geometric schema. The correspondences are investigated between the schema sources to extract features to be integrated in the CityGML database.

An obvious problem when integrating multi source data representing the same area is about geometric and semantic conflicts between competitive data. Indeed, in a preliminary exploration of the two data sets, we have encountered some data discordance due to the mode of representation of the geometry. For instance, an object can be represented in one database and correspond to a group of primitives in the second one (ex: the PICC and IGN buildings are regrouped geometrically according to differing grouping criteria which is respectively the address number and the building function). In general, matching algorithms are based on the distances between geometric locations, the shape of the objects and the topological relations. In our experiment (where the data is isolated), the rules guiding the process of data matching were based on the comparison of some elements (mostly taken together) such as attributes, positions, shapes and geographical names, etc. In some cases, splitting features was necessary to make data matching possible. The strategy of geometric integration was conducted with regards to quality components of source data such as accuracy, resolution, completeness and consistency. Precision and completeness of the 2D geometry were the main criteria. Indeed, the PICC data which is more precise was maintained in most cases expect when completeness and actuality were not good. In other cases, two competitive data can be maintained to populate different CityGML modules. As it was indicated in the CityGML, the grouping concept allows for the aggregation of buildings according to user-defined criteria [9]. Indeed, building groups according to their functions (ex: school building) can be maintained in the thematic module: City Object Group of CityGML. This module can be used to have a generalized city model for some applications that doesn’t require detailed information on buildings. However, a building group with a unique height interpolated according the process of Figure 2 can be unusable for some applications like in military that need intervisibility computations. Figure 4 illustrates a 3D Model in Arcscene obtained by extrusion of building layer resulting from data integration process.

3.3.3.2 3D Reconstruction

Several researchers have studied the issue about 3D reconstruction of buildings and have proposed different methods depending on the initial data source [19–21]. Extruding buildings from footprints is the simplest and the well-known method to construct a 3D building, if the topological relationships between the footprints are not taken into account [22]. Constrained by the nature of data source, we have adopted extrusion as a simplest way to construct 3D buildings. The file resulting from data matching is converted in 3ds format and extruded by the height attribute with FME Workbench program. The 3ds mesh file is then imported in 3dsMax to be structured in different layers: Roof Surface; Ground Surface and Wall Surface for each building.

3.3.3.3 CityGML Conversion

After the 3D reconstruction and building structuring in different boundary surfaces, the buildings are geometrically ready to be converted to CityGML. To perform this task, FME program was used. FME is a Spatial ETL (Extract, Transform and Load) application concept which provides unlimited flexibility in data model transformation, translation and integration [23]. The conversion process was done with several transformers. FME contains over 400 different transformers to carry out different types of restructuring [24]. The challenge was to select the adequate ones for the conversion. Many works have addressed the CityGML conversion from shape file data using FME (see [23]), but there is no unique way to do the conversion because FME provides a large library of transformers and also offers the possibility to develop plug-ins to reconstruct the 3D structure. In this way, the CityGML conversion can be done in one integrated process. Alternatively, we chose to use an external tool (3dsMax) for 3D
Usage, Usability, and Utility of 3D City Models

Figure 4. 3D Buildings after data matching.

modeling which offers an efficient solution to construct complex 3D structures and to prepare data to be converted to CityGML.

After conversion with FME, the resulting writer is a CityGML file which can be viewed via the LandXplorer CityGML viewer program (Figure 5).

3.3.3.4 Data import

For a basic experimentation, we adopted the 3DCityDatabase (version 2.0.6): a free and open source 3D geodatabase to store, represent, and manage virtual 3D city models. The database model, based on CityGML, contains semantically rich, hierarchically structured and multi-scale urban objects facilitating complex GIS modeling and analysis tasks [16]. After creating and configuring an oracle 11g/R2 database instance, the database schema was installed using the creation script (Create_DB.sql) for 3DCityDatabase [25]. The CityGML file resulting from the conversion was successfully imported in the oracle database using the 3D City Database Import/Export Tool (Figure 5): a Java based front-end for the 3D City Database that allows importing and exporting spatial data for a virtual 3D city model. The Oracle SQL Developer is then used to make a connection in the Oracle database and to inspect and query the CityGML tables (Figure 6).
Figure 5. Building model view on LandXplorer CityGML Viewer.

Figure 6. Importing CityGML file using the 3D City Database Import/Export Tool and visualization with Oracle SQL Developer (Surface_Geometry table).
3.4 Discussion

Several methods for acquiring and constructing 3D information from aerial photogrammetry, recognized as the technology more affordable in terms of cost [21], were developed by researchers. They have the particularity to produce the 3D information at the moment of acquisition. Different techniques can be used: capturing detailed geometries and generalizing them by maintaining the link between the detailed geometry and the generalized ones; acquiring data at different scales and using the geometric and semantic matching to build the link between multi scale representations; extracting geometries and establishing the links in the same process [21]. Another method uses parametric modeling of buildings based on predefined models [26].

Our method has the particularity of offering a solution to deal with existing data as a part of the future system. We have suggested a process (geometric in a first level) to integrate data not necessarily acquired to produce a 3D information. Through this experimentation, as a part for data reengineering, we demonstrate that multi source data can be exploited to reach a common model based on CityGML which can be enriched in the future.

4. CONCLUSIONS

In this paper, concepts and main issues about adopting a 3D reference model, as a fundamental component of 3D GIS, are discussed. Collaboration is introduced as a promising approach to face new challenges and adopt a collaborative 3D model in Walloon Region. According to the experiment, existing data can be restructured and converted to CityGML. However, the data reengineering is not trivial. It represents a consistent step to integrate existing data in a collaborative 3D model. Besides of establishing a common 3D model and making agreement on specifications for 3D coproduction, the future challenges are to deal with multi source 3D data integration. The co-produced data will be interoperable but not necessarily with the same quality. Furthermore, other themes like transportation network, terrain and so forth must be addressed to establish rich 3D models to support a large range of applications.

References

European COST Action TU0801


[22] H. Ledoux and M. Metjers. Extruding building footprints to create topologically consistent 3D city models (2009)


