DATA INTEGRATION OF DIFFERENT DOMAINS IN GEO-INFORMATION MANAGEMENT: A RAILWAY INFRASTRUCTURE CASE STUDY

M. Corongiu 1, G. Tucci 1, E. Santoro 1, O. Kourounioti 2

1 Dept of Civil and Environmental Engineering, University of Florence, 50139, via di Santa Marta 3, Firenze, Italy - (manuela.corongiu, grazia.tucci)@unifi.it, enzo.santoro57@gmail.com
2 School of Rural and Surveying Engineering, National Technical University of Athens, 9, Heroon Polytechneiou Street 15780, Athens, Greece - oly_kourounioti@yahoo.gr

Commission IV, WG IV/1

KEY WORDS: 3D city model, multi-dimensional modelling, multi-source data integration, geo-data management, 3D spatial databases, Geographic Information Systems, 3D geo-DBMS, Spatial Data Infrastructure, point clouds, BIM, GIS

ABSTRACT:

A 3D city model is a representation of an urban environment with a three-dimensional geometry of common urban objects and structures, with buildings as the most prominent feature.

In the last decades, 3D city models appear to have been predominantly used for visualisation; however, nowadays they are being increasingly employed in a number of domains and for a broad range of tasks beyond visualisation. The MUIF (Modello Unico dell’Infrastruttura Fisica) project, here illustrated as a case study, refers to the implementation of a single spatial model of the infrastructure of Italy’s railway system (RFI). The authors describe preliminary results and the critical aspects of the study they are carrying out, explaining the processes and methodology to model all datasets into a single integrated spatial model as the reference base for future continuously updates. The case study refers to data collected by different sources and at various resolutions. An integrated spatial Database has been used for modelling topographic 3D objects, traditionally implemented in a 3D city model, as well as other specific 3D objects, related to the railway infrastructure that, usually, aren’t modelled in a 3D city model, following the same methodology as the first ones.

1. INTRODUCTION

Three-dimensional city models have been increasingly applied as data sources or visualization/communication methods in a growing number of fields, such as urban planning and city management, architecture, archaeological reconstruction, tourism, civil engineering, mobile telecommunication, energy supply, navigation, environmental simulation, disaster management, and the game industry (Tack et al., 2012). Semantic modelling of cities requires appropriate qualification of 3D data. In some cases, this can be achieved by an automated process, in others by manual interpretation (Kolbe, 2009).

From the point of view of geomatics, information from 3D modelling has recently achieved a significant level of complexity thanks to the evolution of the technological tools that allow one to survey and manage 3D data in specific models. While most efforts have been addressed to visualization aspects, 3D/4D data management is, however, still not satisfactory within databases. In particular, the Topographic Database (TDB) preserves all data interrogation features, both spatial and alphanumeric, specific to a Geographical Information System. GIS technology, though, cannot yet provide the spatial topological operators for the third dimension, in the same way as those are able to for 2D (Hegenhofer et al., 1994, Clementini et al., 1996). Thus far, the best fitting reference for structuring topological 3D properties has been the B-Rep (Boundary Representation) structure (Vebree et al., 2004). Hence the lack of a standard process to combine and integrate shared 3D datasets and of a way to manage them in a Geo-DBMS (Gröger et al., 2004) is one of the main topics of this paper.

2. STATE OF THE ART

2.1 3D city models

A 3D city model is a representation of an urban environment with a three-dimensional geometry of common urban objects and structures, with buildings as the most prominent feature. The increasing number of applications makes it difficult to keep track of the utilization possibilities of 3D city models; it appears that, despite the near-ubiquity of 3D city models, a comprehensive inventory of 3D applications still does not exist.

Since each 3D application requires its own specific 3D data, a comprehensive inventory can help to link general requirements to specific models. The identification of requirements emerging across domains to generate 3D data has been described in recent studies (Biljecki et al., 2015). In the last decades, 3D city models appear to have been predominantly used for visualisation; however, today they are being increasingly employed in a number of domains and for a broad range of tasks beyond visualisation. In recent years, there have been several initiatives dedicated to semantic topics in 3D city modelling: the Dutch 3D pilot (Goos et al., 2011), the German initiative InGeoForum (Coors et al., 2013), the European COST Action TU0801 (Billen et al., 2015, Billen et al., 2014) and the CityGML standard (Gröger & Plümer, 2012, CityGML, 2012), just to name a few.

CityGML (CityGML, 2012) is an open standardized data model and exchange format to store digital 3D models of cities and landscapes; it defines ways to describe the most common 3D features and objects found in cities (such as buildings, roads, rivers, bridges, vegetation and urban furniture) together with the
relationships between them. It also defines different standard levels of detail (LODs) for the 3D objects, which allows us to represent objects for different applications and purposes. Using CityGML enables us to expand the extensive expertise on 3D city modelling that has accumulated over the years. Today this is a growing field, with a significant body of research and material, with many software tools that support the format, and a long list of potential uses. A wide range of applications exists, such as those for sustainable urban development, water flow modelling, city climate studies, map updating, environmental evaluations, surface analysis, architectural design and 3D visualizations. Although these application fields share a common demand for 3D information, their special requirements with regard to accuracy, detail, actuality and interoperability differ considerably (Tack et al., 2012).

In the past, virtual 3D city models were used mainly for the visualisation or graphical exploration of cityscapes. Nowadays, an increasing number of applications, such as environmental and training simulations, urban planning and facility management, disaster management, homeland security and personal navigation, require additional information about city objects with respect to that given in standardised representations (Kolbe, 2009). A significant example is represented by 3D city modelling applied to transport networks in order to analyse noise from railways or car traffic. For instance, some studies (Lu et al., 2017) simulate traffic noise coming from cars or motorcycles as well as from planes and railroads on a 3D noise map, estimating to what extent this is related to building height. This research draws the conclusion that results on a 3D map are more reliable than those on a 2D grid, as noise has more than two directions in a 3D space. In order to choose the best solution to mitigate the problem, different types of noise in the area as well as different building heights have been considered.

From a topographic point of view, some interesting studies have been conducted in Italy: for instance, in the medieval town of Siena (Figure 1), 3D topographic models have been implemented as a response to the need both to take the third dimension into account and to provide quite detailed scaling (e.g., 1:500 acquisition scale) (Corongiu et al., 2006). The topographic database (TDB) preserves all data interrogation features as well as spatial and alphanumeric analysis features. Technical Cartography from the same TDB has been derived.

The 3D model was conceived as a natural extension of the “volumetric units” borrowed from technical cartography, thanks to the extrusion of each reference surface (Figure 2, Figure 3). This technique allows volume disposition of a building or of any other topographic object to be represented.

![Figure 2 – TDB view of Siena before extrusion.](image1)

![Figure 3 – TDB view of Siena after extrusion.](image2)

Data were organized in a Topographic Database following IntesaGIS National Specifications after their formalization in Italian law (Ministerial Decree of 10 November 2011, 2012). The primary aim of the IntesaGIS project was the development of a Geographical Reference Database: conceived as a shared core with respect to different applications for territory management and supporting different applicative or thematic bases, the database has been used by all concerned government administrations. Moreover, according to TDB Specs, data had to:
- enable the derivation of DB25 (1:25,000 national DB of Italian Military Geographic Institute - IGM) from the database at a larger scale;
- enable cartographic portrayal;
- be in line with international standards and national rules;
- contain metadata information;
- take into account the evolution from 2D to 3D.

However, until now the connection between the implementation of Object-Relational databases (to manage geospatial data infrastructure) and the adoption of a relation-free semantics (which could be used in different applications with 3D models describing spatial objects) is not yet developed out of specific use cases.

### 2.2 Three-dimensional point clouds and BIM-GIS integration

Traditionally, point clouds have been converted to grids, vector objects or other types of data to support further processing in a GIS environment. Lately, though, the use of point clouds has increased through the web, especially for purposes of visualization. In the meantime, the suitability of Database Management Systems (DBMS) for managing point cloud data continues to be debated (van Oosterom et al., 2017).

Today an automatic, fast and cost-efficient construction of 3D city models is still an ongoing research topic. However, recent
progress in 3D geographic data acquisition, data management and 3D visualization have made 3D city models available for a broader range of uses, providing effective solutions for numerous applications (Peters et al., 2017).

In practice, the creation and maintenance of virtual 3D city models is based on a number of independent data sources, since the sustainable management of 3D city models requires close links to existing administrative work flows and databases. A major challenge involves integrating these data sources in a systematic and pragmatic way (Döllner et al., 2006).

Different applications of 3D city models requiring a high degree of photorealistic representation of urban objects, such as for tourism or city planning, already exist. Photorealistic representations increase the quality and quantity of visual information content. In a representation of virtual 3D city models, it is common practice to texture objects, such as building façades or roofs, which have been collected by aerial and ground-based laser scanning in combination with aerial or satellite imagery (Peters et al., 2017).

As laser scanning has become one of the main sources in BIM (Building Information Modelling), the challenges of integrating BIM and GIS into one framework have been the focus of recent research. BIM and GIS can model the built environment in 3D (representing both indoor and outdoor features); additionally, they can be managed in a Database Management System (although for BIM this is usually only done for storage purposes rather than for offering direct query capabilities). They both provide efficient methods for documenting, editing, managing and visualising spatial and non-spatial information, and can represent the world “as is”. They can further model historic data, future planning, modelling outcomes and model data at different scales. BIM has been characterised as “a modelling technology that combines the design and visualisation capabilities of CAD with the rich parametric object and attribute modelling of GIS”. Additionally, for pre-existing buildings or infrastructure, geometry is often generated by laser-scanning (Ellul et al., 2017).

Moreover, the third dimension could be modelled in a relation-free structure, independently from the specific use case application, not only for visualisation scopes, not only for project analysis scopes, not in a specific field of study. For that reason, a single database model could be designed to integrate information coming from different data sources combined in a harmonised way. The challenge of this approach has been not only to solve the 3D data modelling, but also to integrate 3D data coming from different data sources, with different structures or accuracies into one harmonised data model as a base for different objectives: management, manutentions, project, commercial scenarios, etc. For instance, vector data are at 1:1000/1:2000 scales and are supplied from photogrammetry and laser scanner integrated surveys, carried out by plane as well as train along all the railway network and inside stations.

About the availability of 3D spatial operator, the lack of common topological tools both in opensource and commercial software, oriented the control and the management of 3D coordinate specifically, paying attention of the consistency between different objects by a visual or in some cases manual implementation of validation. This aspect impacts on planning resources and the automation processing to model and then implement the geographic infrastructure of interest. In fact, one of the main concern in modelling has been the definition of specific spatial objects related to railways with different functionalities, such as railway switches, intersections and railway artefacts, hereafter called “ASSETS”. Together with all geographical data in the railway neighbourhood, assets have been modelled in an integrated geo-topographic database compliant with Italian national specifications (Ministerial Decree of 10 November 2011, 2012) in order to implement a sharable topographic infrastructure ad a first base geographical reference. A profile and an addition from the national standard has been carried out to implement a system that continuously could be updated thanks to the in situ cartography that Local Administration (e.g. regional or Municipality etc.) timely supply topographic databases their own territory. Regarding to the 4D requirement this approach is only a starting point toward interoperability, but perfectly compliant with the European INSPIRE Directive principles (Directive 2007/2/EC, 2007).

About the availability of 3D spatial operator, the lack of common topological tools both in opensource and commercial software, oriented the control and the management of 3D coordinate specifically, paying attention of the consistency between different objects by a visual or in some cases manual implementation of validation. This aspect impacts on planning resources and the automation processing to model and then implement the geographic infrastructure of interest. In fact, one of the main concern in modelling has been the definition of specific spatial objects related to railways with different functionalities, such as railway switches, intersections and railway artefacts, hereafter called “ASSETS”. Together with all geographical data in the railway neighbourhood, assets have been modelled in an integrated geo-topographic database compliant with Italian national specifications (Ministerial Decree of 10 November 2011, 2012) in order to implement a sharable topographic infrastructure ad a first base geographical reference. A profile and an addition from the national standard has been carried out to implement a system that continuously could be updated thanks to the in situ cartography that Local Administration (e.g. regional or Municipality etc.) timely supply topographic databases their own territory. Regarding to the 4D requirement this approach is only a starting point toward interoperability, but perfectly compliant with the European INSPIRE Directive principles (Directive 2007/2/EC, 2007).

All assets are further related to their respective railway offices in order to be properly managed on site, for this reason a
specific external object-ID have been implemented into the topographic database. Assets together with other datasets (orthophotos, point clouds, DEM, etc.), topographic database and different geographical secondary datasets (cadastral data, geological information etc.) will be made available through a WebGIS interface as well as OGC web services.

3.1 Overview of sources datasets

MUIF survey campaign is carried out using different sources and sensors; hence all data, with their temporal and spatial dimensions must be organised into a single model of spatial railway information.

Source data refers to:
1. an airborne photogrammetric survey (medium Ground Sample Distance – GSD – expected accuracy = 0.08 m);
2. a LiDAR survey (point density expected > = 4 pt/m²) (Figure 4);
3. a Mobile Mapping Survey - MMS (by train). Side by Side images, panoramic images composed of six synchronous shots acquired in six different directions and point clouds of telemetry survey with colours from RGB of panoramic images (Figure 5).

Processed datasets refer to:
1. a railway track centreline;
2. orthophotos (pixel size =0.10 m) (Figure 6);
3. DDTM-DDSM Dense Terrain and Dense Surface Digital Models (resolution =1 m) and a DEM Digital Elevation Model (resolution = 5 m) as derived from LiDAR point clouds;
4. Railway assets and a geo-topographic database at different levels of detail (1:1000 and 1:2000).

All of these datasets were integrated into a single GIS, and some advanced processing analyses were performed in order to evaluate which kind of railway assets should be modelled in 3D. In such a model, information about spatial components (i.e. heights below portals, thicknesses of pylons, etc.) may be semantically structured into a spatial database as DBMS attributes.

3.2 Objective and Processes

The research aimed to adopt different types of datasets in the same GIS environment, making it possible to analyse all aspects of geographic information using spatial operators, for instance to query which infrastructures are shared between transportation by car and by train, or which part of hydrography network impacts with the railway network.

The first step involved re-projecting all datasets toward a single Spatial Reference System (SRS) so that they could be analysed and overlaid together (and not only visualised by a re-projection on the fly). In this way, the combination of data from different sources was achieved.

As the datasets which needed to be combined had different resolutions and came from different sources, processing them into a sole DBMS to achieve integrated content was the main objective. So, information coming from many data sources have been integrated to each class of objects into the database. For instance, the railway network supplied by aerophotogrammetric source is integrated with the point clouds coming of laser by train in underground path (e.g. in a gallery) to obtain a composite track-centerline with an attribute that specify their position relating with the DTM (over ground, in site, underground, etc.).This harmonisation process if almost ordinary for topographic objects, have been more complex for object of specific interest of railway Enterprise, called Asset, that until now have been managed in detailed datasets not integrated with the territorial context and not yet semantically standardised in a topographic 3D model. They have been implemented as 3D spatial components in such a way that a B-Rep could be then derived. This homogenising process is only an intermediate step to obtain a reference database and a semantic model for further spatial 3D object modelling in a geo-DBMS (a work in
The main critical issue was on focus and the related per-
addition, even if their volumes could not be easily defined through basic spatial component surface that can be extruded to obtain a b-Rep visualisation. Then the connection between a complete b-Rep modelling of data and a BIM structure for management will the objective for profiling data.

The point cloud data were separated into two main categories:
- The first consists of the Lidar data and comprises all the area of the railway and its surroundings (buffer zone of about 500 m from the external track), including the ground, buildings, streets, etc. The Lidar data are in LAS 1.2 classification format. There are in total seven classification types, namely, Unassigned, Ground, Low Vegetation, Medium Vegetation, High Vegetation, Building and Noise.
- The second came from a scanner located in the upper part of the diagnostic train for global monitoring of railway infrastructure (a diagnostic system was installed in railway vehicles to perform the monitoring function during their runs). This set consists of 3D points of the railway as well as of building façades, railway ballast and objects close to the tracks. The colours of point clouds were associated according to the RGB of georeferenced panoramic images thanks to the telemetry survey, i.e., the GPS which at every moment estimates the coordinates of the position of both the scanner and the camera.

Correspondence between data of the TDB and the point clouds of the telemetry survey was guaranteed by the IDs of the track centreline. The TDB consists of a number of classes that define spatial objects as a content of a 3D topographic cartography, together with their own spatial components. In addition, properties defining the type of extrusion and the related elevation value were implemented in order to prepare the 3D representation necessary show in their real volumetric position (Figure 7).

Even if all the above datasets were acquired with the same planimetric Spatial Reference System (ETRS2000 RDN2008, EPSG:7794), in cases when the Vertical System referred altitude to ellipsoid instead of orthometric elevation, a geodetic re-projection was carried out.

In order to make orthophotos taken from elevations homogeneous with all other datasets, according to their high-
level resolution (0.10 m pixel size), it was first necessary to produce the TIN of the area. For the Triangulation phase, only Lidar point clouds that had a “ground” classification in the LAS file were selected. Next, the orthophotos were moved to the correct 3D elevation level using the TIN as a reference surface. Afterwards, the Lidar dataset and the TDB buildings were spatially overlaid to classify roof and balcony point clouds (Figure 8).

Since the architectural details of the façades were not easily acquirable from the aerial photogrammetric survey, the integration of façade elements was completed using point clouds coming from the train telemetry survey (Figure 9).

Concerning their density, it was important to visualize point clouds with as many points as possible so that holes could be avoided, and maximum detail obtained. For this reason, the original point clouds were partitioned in order to better manage them and to use the maximum resolution. Obviously, integration with architectonical elements of façades could be attained only for buildings along the railway track.

Specific 3D modelling was performed to focus on railway assets, integrating both aero photogrammetry and train MMS sources. As railway assets represent objects that the Railway Enterprise needs to continuously manage and update for its own institutional purposes, the MUIF project specifically aimed to model them in a GIS. While in traditional technical cartography, assets are represented by points or lines just to mark their position, in the MUIF model assets were related to special information about their volumetric shape (thickness, heights below bridges, etc.). For example, railway portals and pylons have a small footprint, as they pass over the railway and have a specific 3D shape. For this kind of spatial object, not easily viewed in 3D by simple extrusion, a specific integration of models between aero photogrammetry and point clouds was carried out, defining one or more spatial components of each asset class to guarantee a correct 3D view by the extrusion approach (Figure 10).
First of all, the points belonging these objects were isolated from the initial point clouds of the LAS file and digitalised in 3D polygon shapes. Then, each of these polygons was extruded according to its correct height. In this way, these specific objects were included in the TDB as specific classes.

3.3 Critical aspects

3.3.1 Spatial reference System: one of the main issues to take into account in multi-source 3D data integration is the Spatial Reference System (SRS). It should combine different starting points with their own properties and accuracies and move toward a single SRS. What is more, while planimetric interoperable SRS exists, together with its own standard on-the-fly transformation equation, for Vertical Reference Systems (VRS) the lack of global standards requires harmonisation between local geoid to implement orthometric elevations.

3.3.2 Object Semantics: a general 3D model for different thematic domains or for different spatial objects is not available. CityGML is oriented above all to a 3D model of buildings. However, for other kinds of spatial objects lacking homogenous volumes, such as bridges, manufactory objects and transport infrastructure, a more detailed and complex definition of volumes is sometimes needed, as opposed to their simplified boundary representation by planar surfaces. The adoption of a 3D model, then, depends on the thematic context of the application.

3.3.3 Level of Detail: in connection with the spatial object semantic, the level of detail (LoD) required by each domain application must be combined with the accuracy of different data sources. This requirement is a basic constraint in the design of 3D data models.

3.3.4 3D implementation: data integration and modelling must take into account not only realistic 3D visualisation, but above all 3D information content, such that it can be managed and analysed as spatial component information of Spatial DBs in a GIS environment.

4.1 Preliminary results

Three-dimensional datasets are continuously supplied by different sources and for different thematic domains. In the study described above, finding a standard in modelling 3D data, independent of the specific domain, became a requirement to achieve interoperability. At the beginning of the acquisition process, the integration phase was carried out by adopting a single 3D data model as a reference, even though for particular uses a part of the database it is more useful for monitoring aspects as well as other contents answer more performantly to other services such those related to commerce topics, just to cite some requirements. For this reason, main efforts were made to achieve a standard and interoperable approach both among offices of the railway Enterprise and in connection with national standard topographic databases, supplied at large scale (1:1.000 – 1:10.000) by local level Administrations, to combine railway database to each TDB along Italian territory. It was allowed thanks to the implementation of a multidimensional framework from the integration of different datasets in a single spatial database. The purpose was not only to obtain a 3D realistic views, but also to manage the 3D information flow on time, allowing for the future continuously updates coming from multi source data survey So, the design of a 3D data model had to take visualisation as well as information aspects into account, in term of each specific part of the database could be managed by peculiar applications and shared in the reference database for other users.

4.2 Future developments

In light of what has been discussed above, the main aim of the next steps is to define automatic procedures to derive a complete B-Rep of spatial objects, starting from different sources of data. In this context, managing the point cloud survey is especially critical. For future developments, research efforts will be addressed to verify if the point clouds (differently supplied) can be used as a liaison point to be treated following a dual approach: on one hand, the functional life cycle of each single building can be managed, as it happens in a BIM; on the other, all 3D spatial objects can properly be handled in a GIS environment according to their connection at the territorial level. Moreover, a Spatial Data Infrastructures (SDIs) can be a more efficient way of associating different source datasets to semantic territorial objects through their spatial component definitions. A fine-tuned 3D reference model DB will need to be designed independently of specific environmental/urban/topographical domains. The connection between 3D visualisation and 3D DBMS aspects will need to be studied in depth according to 3D semantic data models.

ACKNOWLEDGEMENTS

The authors would like to thank the Italian Railway Network Enterprise (RFI) for making the MUIF datasets available for this case study and would like to thank the e-GEOS company for their support on research activities.

REFERENCES


