

# DATA INTEGRATION FROM DIFFERENT SOURCES TO CREATE 3D VIRTUAL MODEL

A. Erving, P. Rönholm, and M. Nuikka

Institute of Photogrammetry and Remote Sensing, Department of Surveying, Helsinki University of Technology, P.O. Box 1200, FIN-02015 TKK, Finland  
anna.erving@tkk.fi, petri.ronholm@tkk.fi, milka.nuikka@tkk.fi

**KEY WORDS:** 3D model, Photogrammetry, Laser scanning, CityGML, Data integration

## ABSTRACT:

3D modelling of virtual environments is a fast developing field of research, and the need for virtual models is increasing rapidly. This paper presents a work, where the aim was to test virtual 3D model production phases, when there are several different data sources. Photogrammetric and laser scanning techniques are utilized in data acquisition. In addition, other available sources, like a base map, have been used to obtain an extensive data set for different modelling purposes. The use of different data sources brings along many challenges. For example, multiple data formats, various accuracies, coordinate systems and level of details must be taken into consideration. In this work, LandXplorer Studio Professional software is used in data integration and creation of the final model. The software is designed for the production of models, which can be directly converted into CityGML data format. CityGML is a standardized and semantic information model designed for storage and exchange of virtual 3D city models. In this research, the final model of the experimental work consists of ground surface, road, road structures like traffic lights and traffic signs, a building and vegetation. It was noticed that the 3D virtual model construction process is often complex and includes various phases when different data sets are integrated.

## 1. INTRODUCTION

The use of three dimensional virtual models has expanded into many application fields. It has been popular to model city centres or other urban areas including buildings, roads, road structures like traffic lights and traffic signs, and vegetation. Also, remarkably large urban areas have been modelled during the recent years, like the city of Berlin including building models for an area of 250 km<sup>2</sup> (Döllner et al., 2006). Some of the modelled city centres are accessible through Google Earth.

Normally, a model consisting of different objects has to be created by using data sets from several sources and with several data formats. In that case, one has to manage merging this diverse material into a uniform representation of the environments.

In this research, an urban area has been modelled by using multiple data sources. As the virtual model is constructed from different data sets, there is a great challenge to adapt them all to work together. One has to pay attention to various file formats, accuracies, coordinate systems and level of details. If these are not considered during the model construction process, it may, among others, affect to the geometrical accuracy and visual appearance of the final product.

Most of the data used in this work is geometrically accurate, based on measurements. Once we have a proper basic material, also the model with geometrical properties can be successfully produced. Furthermore, a trend of adding semantic and topological properties into models is now becoming a desired way to build virtual models. For a simple flooding simulation of an urban area, a normal geometrical model can be sufficient. However, if we additionally want to know statistics concerning large areas, such as what kind of objects suffer from

the flooding, we need to know more about the model (Schulte and Coors, 2008). This kind of information could be, for example, heights of buildings on that area and the amount of people living around. Semantic and topological data enable many interesting applications such as spatial data mining, thematic queries, semantic-supported navigation or augmented reality, just to name few. For example, Döllner et al. (2005) described the use of virtual models in city planning. Benefits can be directed in model production phase, as well. When integrating different model parts together, the knowledge of the semantic and topological properties can help to build more uniform model (Stadler et al., 2007).

The aim of the work was to explore the possibilities of using different data sets and integrate them into a one three-dimensional (3D) virtual model. The interest was in the workflow and the problems faced during the work. It was studied, how individual data sets should be processed, before they could be used together as a complete model.

Laser scanning point clouds, images, and vector data from base map were used as a source data for completing the model. Because of the different coordinate systems, file formats and other properties, the material had to be processed individually before integrating the data sets together. The merging phase was performed with LandXplorer Studio Professional software (Autodesk, 2008), which also had its own restrictions related to data handling.

During the work, many problems were faced when merging multi-source data sets. Additionally, a new standard and data model, City Geography Markup Language (CityGML), was adopted into model construction. LandXplorer software is designed for the production of models, which can be directly

converted into CityGML models by exporting them into the correct file format. As a result, several different data sources were used in creation of virtual 3D model of a small test area in Espoo, Finland.

## 2. CITYGML

CityGML is an XML-based syntax of a data model (Figure 1), and it has now been adopted as an official standard by OGC (Open Geospatial Consortium). It has been designed to be an open data model, which could be commonly used for storage and exchange of virtual 3D city models. In addition to the objects geometrical properties, CityGML model defines the semantical, topological and appearance features of the object model. For example, semantic information of a building could be its year of construction, the number of floors etc. The possibility to add semantic data to a 3D model increases the usage of applications compared to a normal geometrical model. More complete descriptions about CityGML can be found in Kolbe et al. (2006), Erving (2007) and Gröger et al. (2008), for example.

```

<cityObjectMember>
  <Building>
    <stringAttribute name="_name">
      <value>main_building</value>
    </stringAttribute>
    <stringAttribute name="yearOfConstruction">
      <value>1990</value>
    </stringAttribute>
    <intAttribute name="measuredHeight">
      <value>17</value>
    </intAttribute>
    <intAttribute name="storeysAboveGround">
      <value>5</value>
    </intAttribute>
    <intAttribute name="storeysBelowGround">
      <value>1</value>
    </intAttribute>
    ...

  <lod3MultiSurface>
    <gml:MultiSurface>
      <gml:surfaceMember>
        <gml:Polygon gml:id="tempid6570">
          <gml:exterior>
            <gml:LinearRing>
              <gml:posList>2545291.0569998      6675108.80525163
9.50705      2545285.02699995      6675107.92024949      9.50705
2545285.08500004      6675107.64525178      9.50705      2545291.0569998
6675108.80525163      9.50705 </gml:posList>
            </gml:LinearRing>
          </gml:exterior>
        </gml:Polygon>
      </gml:surfaceMember>
    </gml:MultiSurface>
  </lod3MultiSurface>

```

Figure 1: A code example of a CityGML information model, when the 3D model has been exported from Landxplorer software to GML format.

CityGML allows many tools for visualization. The data model supports the level of detail (LOD) representation. The model can consist of five different levels of details, which help to present different object in a large 3D model. For buildings the levels are block-models (LOD-1), geometry-models (LOD-2), architectural models (LOD-3), and detailed indoor models

(LOD-4). In addition, LOD-0 can present elevation models in 2.5D. Using multiple LODs in one model intensifies and clarifies the visualization task. As a rule of thumb, the further away the object is zoomed out, the lower the level of detail needs to be. On the other hand, different levels can be used also for highlighting places of interest. For example, the most important buildings of city centres can be visualized with higher level of detail than the more common ones.

Our final model can be converted into CityGML data format with LandXplorer, which enables to include semantic information to the model or to a sub-part of the model. Since CityGML is a standardized data model, it will hopefully decrease the handling of different presentations of 3D models in the future. Other software capable to handle CityGML data format on some level are Aristoteles Viewer (<http://www.ikg.uni-bonn.de/aristoteles/index.php/Aristoteles>, accessed 27.1.2009) and the products of Snowflake Software (<http://www.snowflakesoftware.co.uk/products/>, accessed 27.1.2009).

CityGML has a great potential to become a world-wide standard for 3D city models and thereby solve many of current difficulties to exchange data between different software having their own data formats. At least many similar markup languages, such as HyperText Markup Language (HTML) and eXtensible Markup Language (XML), have proven to be tremendous success. Similar success for CityGML requires software developers to realize advantages of common data format.

## 3. MATERIALS

The test site of the experimental work in this research was located in Espoo, Finland. Digital surface model (DSM) of the area of interest was created from airborne laser scanning (ALS) point cloud (Figure 2). The data was collected in 2002 using TopEye MK 1 implementation. The point cloud density was 2.3 points per m<sup>2</sup> in average, flying height was 200 m, laser beam size was approximately 20 cm, the scanning angle was +20°, and the strip width was 110 m.

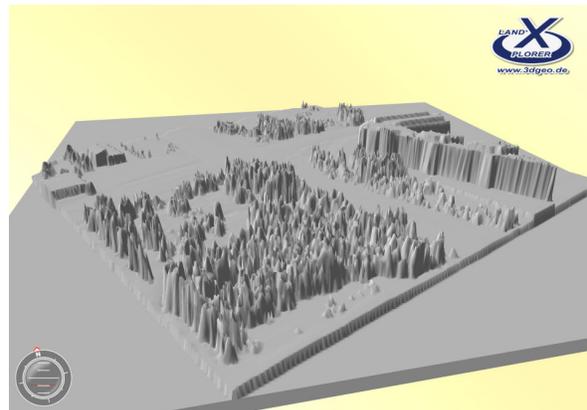


Figure 2. DSM created using airborne laser scanning.

Images from Pictometry application were utilized in DSM texturing. In addition, images were used in finding dimensions of one block building and providing façade textures for it (Pictometry, Blom Kartta Oy, Finland). Pictometry images included ortho photos and oblique aerial images. Ortho photos

were used as a texture on top of a digital surface model. Oblique aerial images were used as source of façade textures in block building model.

3D model constructed from both aerial and terrestrial images was used to visualize a road crossing more detailed in the final model (Figure 3). In this case, images were not from Pictometry, but from another aerial imagery. Calibration parameters of the used cameras were known (Nikon D100, Sony DSC-V1, and Hasselblad). 3D model of the road crossing included the road pavement, traffic lights, and traffic signs, with photorealistic textures. The modelling was done using iWitness software (Fraser and Hanley, 2004), which is specialized in photogrammetric modelling from convergent photography.

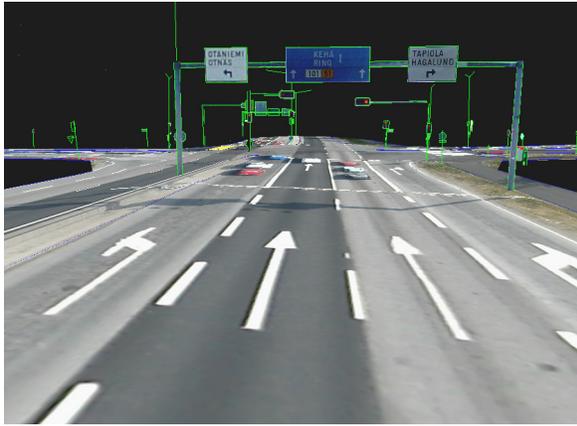


Figure 3. Photorealistic 3D model of a road crossing. (Milka Nuikka, 2007)

In addition to images and laser scanning data, also a base map from the City of Espoo was used in block building construction. The base map consisted of vector data, such as contour lines and top plans of the buildings and roads (Figure 4).

#### 4. DATA PROCESSING

The aim was to build a photorealistic 3D model using the data sets listed above. Before that, different data sets were processed with suitable software in each case.

ALS-based digital surface model was edited by removing points representing one building. Instead, a block building model of this object was constructed based on vector data of a base map.

Façade textures of block building were taken from oblique aerial images. Images were pre-processed with Photoshop in order to reduce the perspective distortions in images and to crop the façades out of the background. Processing of the images is not available in LandXplorer.

Photogrammetric 3D model was measured using iWitness modelling software, in which the only possible export format with textures is VRML (Virtual Reality Modelling Language). On the other hand, LandXplorer doesn't accept VRML as an import file format. Therefore, the format transformation had to be carried out with a third application. In addition, the coordinate axes of a photogrammetric 3D model were divergent compared to the global coordinate system in LandXplorer.

Hence, the model had to be transformed before it was ready to be imported into LandXplorer.

Vector data appeared to be in a different coordinate system compared to the other data sets in virtual model. Therefore, a transformation was required between these two coordinate systems. After the transformation, all data in virtual model was in the same coordinate system (Figure 4). For getting the vector data prepared for importing to LandXplorer, it also needed some topological changes. Originally, the top plan of a complex building consisted of individual lines. For LandXplorer, however, a closed polygon was needed (Figure 5).

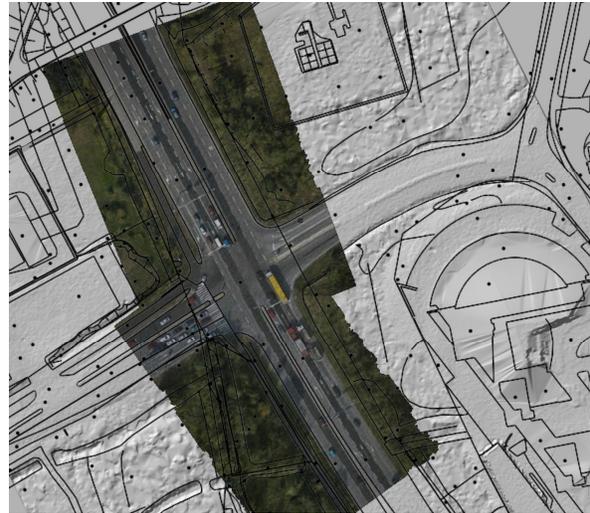


Figure 4. Vector data and ortho photo on the top of the elevation model. All data sets were transformed in the same coordinate system.

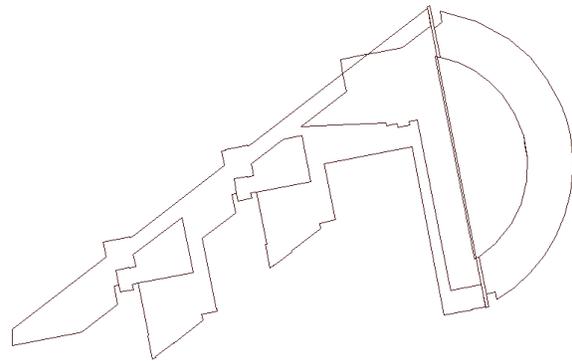


Figure 5. Vector data from the base map presenting a footprint of a complex building.

#### 5. DATA INTEGRATION

In this work, LandXplorer Studio Professional from Autodesk has been utilized to combine individual datasets in order to create a virtual 3D model of an urban area. LandXplorer Studio Professional is software developed for merging data sets and visualization of 3D models. With capability to attach semantic

data to objects and convert the model into CityGML information model, LandXplorer appears to be a useful tool for data integration and CityGML model construction. In addition, objects or parts of objects can have references to external databases, which can provide more semantic information to the model.

After all the data was pre-processed for integration, they were imported into LandXplorer Studio Professional. In LandXplorer, some final tuning between data sets was still possible. For example, the orientation parameters for Pictometry ortho photos were only approximate, and they had to be visually aligned with the surface model. The orientation of ortho photo was realized using a separate world file, which defines parameters for x- and y-coordinates of the upper left corner of the image. In addition, the pixel size of the image was entered to the file. If there was a need to change the position of the photo, it was done by changing the file values.

Block building was visualized in LandXplorer based on vector data. Because vector data from the base map was edited to be as closed polygons, they could be automatically raised into 3D blocks with a defined height value. If only individual lines existed, separate walls would have appeared instead of a solid object. Due to the software functionalities, the bottom of the block building lay on the sea level, where the height value was zero in the used coordinate system (Figure 6).

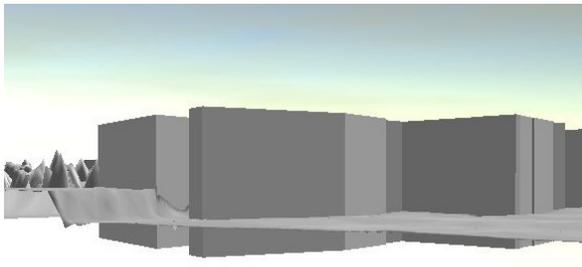


Figure 6. Incorrect elevation of a block building was clearly visible.

As part of the object now occurred under the terrain model, it had to be raised to the ground surface. This was done with the software's function "Drop buildings on terrain". As the height of the building was correct all the time, the shift was done only in z direction with the fixed object (Figure 7).

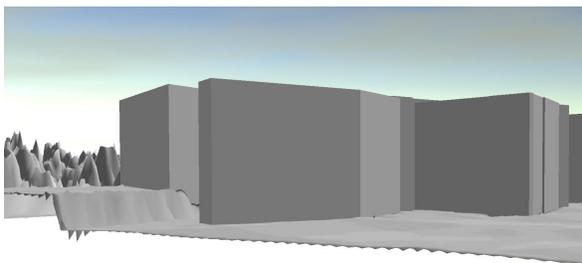


Figure 7. Correct elevation of a block building after raising the object.

The block building height values were obtained using Electronic Field Study software, which is designed to deal with

Pictometry data. As there existed a coarse elevation model behind the images, it was possible to measure height values for buildings using monopleting (Figure 8). In this manner, approximate height values were obtained for block buildings. Because the building was quite complex, it was also necessary that the vector data polygons were formed by separating different parts of the building having different heights.



Figure 8. Obtaining height values from Pictometry application.

The textures of façades for the block buildings were acquired from oblique aerial images, as it was mentioned in Chapter 4. In principle, every plane on the façade differed from others. Therefore, each plane was associated with corresponding unique texture file. In some cases due to occlusions, however, it was not possible to add original texture, and therefore some duplicates were used.

The 3D model of the road crossing fitted well to the virtual model, when it was imported into LandXplorer. The cars appeared flat on the road surface, because they were visible on the aerial image texture but not in the DSM (Figure 9).



Figure 9. 3D model of the road crossing imported into virtual model.

Illumination conditions, perspective differences and reflections may change colours of textures. For virtual models it would be beneficial, if textures were homogenous in respect of these phenomena. For example, natural shadows in textures may

disturb artificial illumination of virtual 3D model. In our case, however, the images were not enhanced radiometrically.

## 6. RESULTS

CityGML is one option for a better compatibility between 3D models and other data sets. It's open source, meaning that anyone can use it without a charge. Among many others, photogrammetric data and laser scanning point clouds can be used as raw material.

LandXplorer Studio Professional proved to be a useful tool for merging various data sets. It has been designed to work with CityGML information models. It supports many different file formats, and can handle surprisingly large data sets, especially in the case of terrain models.

As the final product can be converted to a CityGML data model, it can also be used in semantic queries. This will be useful when dealing with large data sets, like city models, since it is possible to search different objects from the model based on given parameters.

The final model consists of a surface model, 3D model of the road crossing, and one complex block building (Figure 10). The workflow for creation of CityGML model from individual data sets is presented in Figure 11.

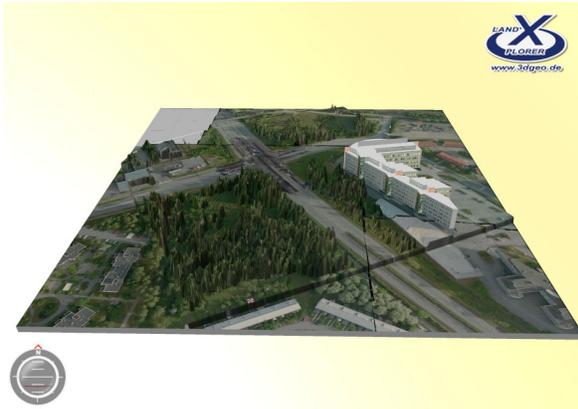


Figure 10. The final photorealistic virtual model.

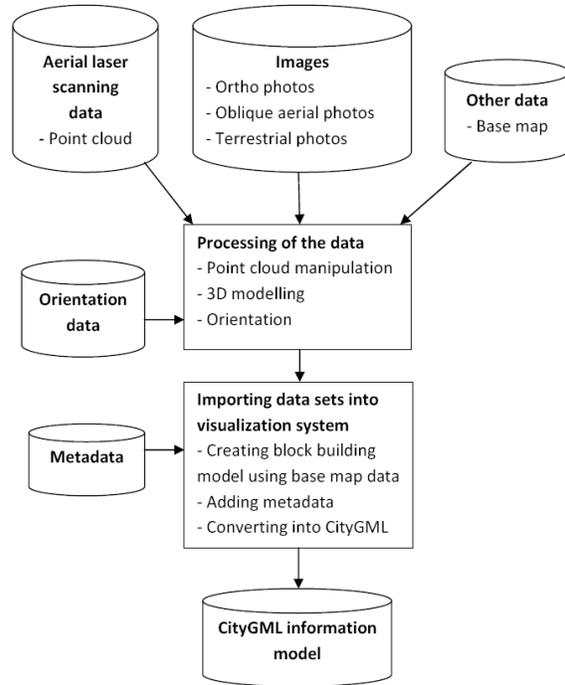


Figure 11: Data sources and data processing steps towards CityGML information model.

## 7. DISCUSSION

During the work, several problems with data integration were noticed. Some of them are naturally caused by the software in use. Software utilizing CityGML are still at the early stage and it is not yet possible to import data from all available data formats. Specially, if data acquisition software uses non-standard data formats, it may be difficult to convert data to another format without losing some information. In practise, successful data conversion may require the use of several software.

One challenging part is the coverage of the data. Missing data may cause holes in the resulting models, especially in the case of textures. If the model textures are incomplete, one option is to copy one texture into several places, as it was done in this case as well. If the textures are not completely real ones, however, the realistic appearance does not necessarily realize across the whole model.

Also, the quality of image orientations affects a lot to the results. If sufficient orientation parameters are available, texture mapping can be done to match the reality. In this case, the façade textures were processed and rectified roughly and manually using Photoshop. This step would need a better solution and more automatic processing in order to be efficient.

Oblique aerial images proved to be well suitable for façade texturing. Although the images are photographed from the air, the façades are quite well visible. Pictometry image acquisition method provides a lot of images, which makes it easy to find good viewing perspectives.

Alternative source of textures to oblique aerial images are terrestrial images. However, the data acquisition of terrestrial images may be time consuming. Therefore, mobile mapping systems that utilize GPS and inertial data may intensify the data collection process.

In many cases, the block models provide sufficient level of detail for visualization, especially when producing large city models. Block models are lighter to handle and faster to produce than more detailed models. The basic data for block models, like base map or cadastre map, is available in many cases. More challenging task may be to get accurately the height value for each building or building part, if such level of detail is needed.

Texture mapping will give the model an appearance that the amount of details would be much more than it is in reality. However, the details coming from image textures are usually not geometrically accurate, since, for example, the block model has certain restrictions when it comes to the shape of the object.

Pre-processing of each data set before integration and final model production may be time-consuming. Typically, data needs adjustments before it can be merged together successfully. LandXplorer, for example, does not yet have all data editing opportunities, which are needed.

Automatic methods for 3D model production are not yet fully functional. The development work is an ongoing process in many research projects these days. The diversity of the modelled objects makes the full automation very challenging. However, if the model production process could be fully automated, the processing of large data sets would be very efficient.

In block building production, automatic methods are already available (Ross et al., 2006). The only phase, which needs more manual attention, is the texture mapping. Texture mapping could be done fully automatic, but the quality of the textures will suffer. Problems occur, when there are occluding objects on the image, which will be attached to the textured surface.

Automation is difficult also with detailed photorealistic 3D modelling tasks. However, terrestrial laser scanning is one way to achieve highly detailed object models relatively quickly. Even using laser scanning, the processing of the data still requires lots of time and also manual inspection.

Required level of details define, how automatic process can be obtained. Also, the available data and its properties have an important role. In many cases, semi-automatic methods are utilized. Kaartinen et al. (2005) have compared different building extraction methods and the accuracies of the resulting models.

## 8. CONCLUSION

In this paper, the virtual model of an urban area was created. LandXplorer Studio Professional software was used as a primary visualization environment. The software proved to be usable to integrate 3D data from different data sources. It is also capable to convert the models into CityGML information model.

Digital elevation model was created using airborne laser scanning data. Triangulated ground surface was textured with ortho images. On the top of textured ground surface, vectorized objects, such as image-derived 3D model of the road crossing and textured block model, were visualized.

The challenge of using data from different sources in one virtual model consists e.g. various coordinate systems, data contents, file formats, level of details, and accuracies. These aspects should be considered before model construction. Additionally, sufficient visual appearance is needed in certain applications. In many cases, the model construction process is complex and includes various phases, when different data sets will be integrated. Finally, the usability of the 3D model for further use should be ensured.

With an experimental work it was noticed, that the 3D virtual model construction process is a challenging task, when multiple data sources will be merged. CityGML appears to be a promising format for visualization, storage and exporting 3D virtual models of urban areas. The success of CityGML, however, will be dependent on software developers, whether they are willing to adopt the format or not.

## REFERENCES

- Autodesk (2008): LandXplorer Studio Professional Software System, [www.3dgeo.de](http://www.3dgeo.de).
- Döllner, J., Kolbe, T. H., Liecke, F., Sgouros, T. and Teichmann, K., 2006. The Virtual 3D City Model of Berlin - Managing, Integrating and Communicating Complex Urban Information, 25th International Symposium on Urban Data Management UDMS 2006 in Aalborg, Denmark.
- Erving, A., 2007. CityGML mallinnus fotogrammetristei. Maanmittaustieteiden päivät 2007, Espoo, Maanmittaustieteiden seura ry, p. 34-40 (in Finnish).
- Fraser, C.S. and Hanley, H.B., 2004. Developments in Close-Range Photogrammetry for 3D Modelling: the iWitness Example. International Workshop: Processing and Visualization using High-Resolution Imagery, Pitsanulok, Thailand, 18-20 Nov., 4p. (on CD-ROM).
- Gröger G., Kolbe, T., Czerwinski, A. and Nagel, C. (editors), 2008. OpenGIS Consortium, Inc. OpenGIS City Geography Markup Language (CityGML) Encoding Standard, version 1.0.0, [http://portal.opengeospatial.org/files/?artifact\\_id=28802](http://portal.opengeospatial.org/files/?artifact_id=28802). (accessed 14.1.2009)
- Kaartinen, H., Hyypä, J., Gülch, E., Vosselman, G., Hyypä, H., Matikainen, L., Hofmann, A. D., Mäder, U., Persson, Å., Söderman, U., Elmqvist, M., Ruiz, A., Dragoja, M., Flamanc, D., Maillet, G., Kersten, T., Carl, J., Hau, R., Wild, E., Frederiksen, L., Holmgaard, J. and Vester, K., 2005. Accuracy of 3D city models: EuroSDR comparison, ISPRS WG III/3, III/4, V/3 Workshop.
- Kolbe, T. H., Gröger, G. and Plümer, L., 2005. CityGML – Interoperable Access to 3D City Models, Proceedings of the Int. Symposium on Geo-information for Disaster Management on 21.-23. March 2005, Delft.

Ross, L., Kleinschmit, B., Döllner, J. and Buchholz, H., 2006. Automated Transformation of 2D Vector-Based Plans to 3D Geovirtual Environments, In Buhmann, E.; Ervin, S.; Jorgeson, I. & Strobl, J. (ed.) Trends in Knowledge-Based Landscape Modeling - Proceedings at Anhalt University of Applied Sciences., Heidelberg (Wichmann), 2006, 138-144.

Schulte, C. and Coors, V., 2008, Development of a CityGML ADE for dynamic 3D flood information, Joint ISCRAM-CHINA and GI4DM Conference on Information Systems for Crisis Management, August 4 – 6, 2008, Harbin, China.

Stadler, A. and Kolbe, T. H., 2007. Spatio-semantic coherence in the integration of 3D city models, 5th International ISPRS Symposium on Spatial Quality ISSDQ 2007 in Enschede, The Netherlands.

#### **ACKNOWLEDGEMENTS**

This study was supported by Ministry of Agriculture and Forestry in Finland with decision number 31155 and Academy of Finland with decision numbers 124574 and 121964. Blom Kartta Oy in Finland has provided us Pictometry test data and the Electronic Field Survey -software.