# Impact of the quality of spatial 3D city models on sensor networks placement optimization

Meysam Argany[1], Mir Abolfazl Mostafavi[1], Vahab Akbarzadeh[2], Christian Gagné[2], Reda Yaagoubi[3]

[1]Center for Research in Geomatics, Department of Geomatics, Université Laval, Quebec, Canada
meysam.argany.1@ulaval.ca
mir-abolfazl.mostafavi@scg.ulaval.ca
[2]Computer Vision and System Laboratory, Department of Electrical Engineering and Computer Engineering, Université Laval, Quebec, Canada
vahab.akbarzadeh.1@ulaval.ca
christian.gagne@gel.ulaval.ca
[3]Department of Geomatics, Faculty of Environmental Design, King Abdulaziz University, Jeddah, Saudi Arabia
ryaagoubi@kau.edu.sa

***Abstract.***—Sensor networks are increasingly used for tracking, monitoring and observing spatial dynamic phenomena in the real world (e.g. urban area). In order to ensure an efficient deployment of a sensor network, several optimization algorithms have been proposed in recent years. Most of these algorithms often rely on oversimplified sensor models. In addition, they do not consider information on the terrain topography, city models, and the presence of diverse obstacles in the sensing area (e.g. buildings, trees, poles). Only some of those optimization algorithms attempt to consider the terrain information in the optimization of a sensor network deployment. However, most of these algorithms consider that the spatial models used for this purpose are perfect representations of the reality and are not sensitive to the quality of the information. However, spatial models are simplified representations of a complex reality, and hence are inherently uncertain. In this paper we will investigate the impact of the spatial data quality on the optimization of a sensor network and its spatial coverage in an urban area. For this purpose, we will investigate specific implications of spatial data quality criteria for a 3D city model that will be used in sensor network optimization algorithms. Then, we will analyze the impact of some of those criteria on the estimation of sensor network coverage. Afterwards, a case study for sensor network deployment in an urban area will be presented. This case study will demonstrate the impact of 3D city models quality on the estimation of coverage using global and local optimization algorithms. Finally, the results obtained from this experimentation will be presented and discussed.

## 1- Introduction

Sensor networks are increasingly used for tracking, monitoring and observing spatial dynamic phenomena of the real world (Nittel 2009). The benefit of using such networks is to access remote or harsh areas and observe phenomena in these locations at the lowest cost as possible. The cost of a sensor network deployment depends mainly on the number of sensors used and how these sensors are placed in the environment to be monitored. Hence, in order to maximize the spatial coverage of such networks, optimization algorithms can be used to find the best position for each sensor in the network. However, most of the proposed placement algorithms do not consider the nature of real environments (Aziz et al. 2009, Bharathidasan et al. 2003, Nittel 2009). In addition, the few works that take into account the environmental information in their methods (Wang and Tseng 2008, Akbarzadeh et al. 2011) do not study the impact of the inherent uncertainty of spatial data in the estimation of sensor network coverage.

There are many objects and obstacles in the environment that may constrain the spatial coverage of a sensor network. Therefore, it is necessary to consider these elements in sensor network optimization algorithms. For example, in an urban area, the presence of buildings, roads, streets, trees, and poles should be considered in sensor deployment. In a natural area, the topography of the terrain and other properties of the environment such as vegetation must be known. Spatial models are very rich sources of geospatial information that can be used inside the optimization algorithms. However, spatial models are simplified representations of a complex reality, and hence are inherently uncertain. The uncertainty in spatial data may be related to the methods used for the acquisition, processing, or manipulating of spatial data, and it may significantly affect the spatial coverage of a sensor network.

Some of the most important types of datasets, which are used as spatial models in sensor network deployment, are Digital Terrain Models (DTM) and Digital Surface Models (DSM). The quality of these models is varied and depends on the accuracy of the initial datasets, which are used to produce them as well as the instruments, which have been used to collect those datasets. For example, both digital terrain models and digital surface models may have some inaccuracies which involve some unintentional errors in final results. Since we have errors and inaccuracies within the initial datasets, it is inevitable that these errors will be propagated when these dataset are used for deployment of sensor networks. So, accuracy of sensor placement strongly depends on the quality of spatial models that will be used in optimization algorithms. Also, the communication between sensors in a given network may be affected by the quality of the data as well. In fact, the position of sensors and their communication range are important to ensure reliable communication between sensors.

In this paper, we investigate the data quality elements with an emphasis on those that are the most relevant for 3D city models. We will study the impact of those elements on sensor network coverage estimation. Then, we will investigate the impact of the 3D dataset's quality, which will be introduced as initial input in the sensor network deployment optimization algorithms on the final results. Our goal is to determine how sensitive different optimization algorithms are to the quality of input datasets and what their behaviour will be.

The remainder of this paper is organized as follows. Section 2 presents a literature review describing various models and solutions of the sensor deployment optimization based on 3D city models. Local and global approaches for sensor deployment optimization are discussed in this section. In Section 3, the quality elements for 3D city models are introduced. First, standard spatial data quality elements are presented and then most relevant data quality elements for 3D city models are further investigated and their implications for sensor placement are discussed. Section 4 presents an analysis of the quality impact of 3D city models on the sensor network deployment. The issue of how 3D models quality affects the results of optimization methods will be discussed. Section 5 contains the experimentations and results. Several maps with different quality levels have been prepared and tested with three optimization algorithms. The sensitivity of the optimization methods to the quality of input data is investigated in that section. Finally, Section 6 concludes the paper with discussion of the results and proposal of new avenues for future work.

## 2- Sensor Network Deployment Optimization Based on 3D City Models

Efficient sensor network deployment is an important issue in the sensor network field, as it affects the coverage and communication between sensors in the network. Nodes use their sensing modules to detect events occurring in the region of interest (e.g. urban area). Each sensor is assumed to have a sensing range, which may be constrained by the phenomenon being sensed as well as the environmental conditions. Hence, obstacles and environmental conditions affect network coverage and may result in holes in the sensing area. The definition of coverage differs from one application to another (Aziz et al. 2009, Ahmed et al. 2005, Ghosh et al. 2008, Huang et al. 2005, Megerian et al. 2005, Meguerdichian et al. 2001). In this study, the definition of coverage is based on a direct visibility between the observer and the target point (e.g., camera for traffic monitoring) (Fig. 1). The coverage of a point in a sensor network means that this point is located in the sensing range of at least one sensor node. The coverage area of each node is usually assumed to be uniform in all directions. In this case, the sensing range is represented by a disk around the sensor. Failing this condition for some points in the region of interest will result in coverage holes (Ahmed et al 2005).


**Fig. 1.** Direct visibility between an observer and a target, point A is visible while point C is invisible because its line-of-sight is concealed at point B.

Hence, one important issue in sensor network deployment is finding the best sensor position to cover the region of interest. Regarding the mentioned definition of coverage in sensor network, the coverage problem basically means placing a minimum number of nodes in an environment, such that every point in the sensing field is optimally covered (Aziz et al 2009, Ghosh and Das 2008). Nodes can either be placed manually at predetermined locations or dropped randomly in the environment and then are repositioned to their optimal locations. It is difficult to find a random scattering solution that satisfies all the coverage and communication conditions between sensors.

There are several approaches in the literature to solve the problem of sensor network coverage (Niewiadomska et al. 2009, Romoozi et al. 2010, Ghosh et al. 2008). In general, these approaches are classified into global and local optimization approaches. Global optimization approaches are used to find the global optima of a function (coverage function) or a set of functions for the whole study area. Conversely, local optimization methods are used to find local optima among a number of candidate solutions. Candidate solutions here could be the sensors positions or final coverage, which is supposed to be optimized according to the coverage function. Local methods start with an initial value in the space of candidate solutions and then iteratively move to neighbour values or solutions by applying local changes until the optimal solution is found or a time bound is achieved.

## 2.1 Global optimization approaches

Simulated Annealing (SA) and Covariance Matrix Adaption Evolutionary Strategy (CMA-ES) are two examples of global optimization methods used for sensor network deployment (Akbarzadeh et al. 2010, 2011). These methods will be used in Section 5 to compare the impact of data quality on sensor network deployment given their performance and popularity in global optimizations (Akbarzadeh et al. 2011).

Simulated Annealing (Kirkpatrick et al. 1983) is a classical metaheuristic optimization algorithm, which is inspired by the annealing process of material in metallurgy. In fact, temperature is the controlling mechanism used to convert material from a high-energy state into a low-energy solid condition. This process is imitated in SA, where the temperature controls the number and spread of accessible solutions from a given solution in the search space. SA starts with random sensor positions in the 3D study area with a high initial temperature to allow a random walk in the search space. As the temperature is gradually decreasing the system becomes greedier, only to allow moves in the search space which improve the performance of the solution to find optimized positions which best served coverage. The process is completed with a temperature close to zero. To calculate the coverage, a coverage function, which will be introduced in Section 2.3, is supposed to be optimized by means of an optimization algorithm.

CMA-ES is part of the evolutionary algorithm family. It is a black-box stochastic optimization method, in which new candidate solutions (sensors positions) are sampled according to a multivariate Gaussian distribution, which is adapted in the course of the optimization (Hansen and Ostermeier 2001). For sensor network deployment optimization, the initial position and orientation of sensors in a 3D model can be considered as a candidate solution. So, any variations or inaccuracies in the 3D model affect the position and orientation of the sensors and hence, directly impact the formation of the next solutions. The sensor positions will be evolved through the optimization and finally, the solution with the best coverage is selected as the final result (Akbarzadeh et al. 2011).

## 2.2 Local optimization approaches

The second category of optimization algorithms for sensor network deployment is the local approaches. Some geometric solutions found in the literature are taking into account the spatial relations between the elements of 3D model (search space). When there is not enough information available about the environment, sensors are deployed randomly at the first placement, and then some deployment strategies take advantage of mobility and try to relocate sensors from their initial position to optimize the network coverage. In these cases, spatial information, sensor's positions and movement strategies will be provided based on 3D models. VECtor-based and VORonoi-based algorithms are two mobility-based methods that use Voronoi diagram in their approaches (Argany et al. 2011). The spatial coverage of sensor networks in 3D models is much related to the spatial distribution of the sensors. In other words, the geometric solutions try to distribute the sensors in the environment by using 3D models so that as much coverage as possible will be obtained. These approaches can be used to detect coverage holes in 3D datasets as well as healing those holes.

The VORonoi-based algorithm (VOR) is a pulling strategy; this means that sensors cover their local maximum holes. This method has been selected to study the impact of data quality in Section 5 because of its geometrical performance and ability to model the environment (Argany et al. 2011). In this algorithm, each sensor moves toward its farthest Voronoi vertex until this vertex is covered. The disadvantage of the VOR algorithm is that each sensor may be selected to move but there is no criterion to define where it should stop moving. A 3D model of the environment can help us to define this threshold, which means that sensors stop moving when they arrive at the point with a higher elevation than their initial position. The line of movement corresponds to the line between the initial sensor location and farthest Voronoi vertex. In the rest of the paper we call this approach the enhanced VOR algorithm. Here, Voronoi cells define the regions in the study area which should be covered by the sensor inside the cell. Since the sensing range of sensors is limited, then some holes may exist beyond the sensing area of the sensors (Fig 2).

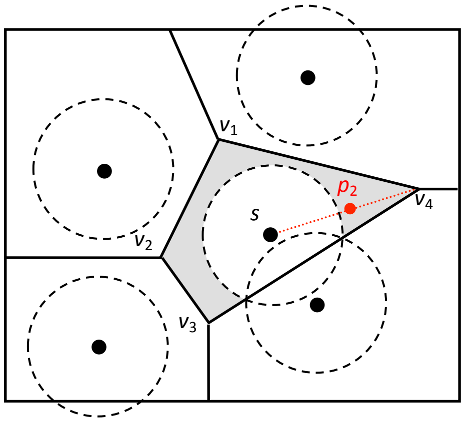



**Fig. 2.** Movement of a sensor in the VOR algorithm.

## 2.3 Coverage estimation using 3D city models

In this paper, coverage is defined based on the concept of line-of-sight. Line-of-sight can be defined as the direct visibility between an observer and a target point. Given the sensor position $p_i$, if there is no obstacle between the $p_i$ and the target point $\boldsymbol{q}$, then the latter is visible. Also, if $\boldsymbol{q}$ is in the sensing range of $s_i$, the coverage is achieved (Fig. 1). Viewshed is another term which is used in optimization algorithms. It is defined as an area in the maps that is visible from a specific sensor position. Viewshed algorithms use elevation of each cell in the DTM to determine visibility to or from a particular cell (sensor positions). The visibility depends on the following notions: observation points, horizontal and elevation coordinates $(x_i, y_i, z_i)$, vertical offsets (the vertical distance to be added to the $z$ coordinate value of a location on the surface), horizontal and vertical sensor orientation $(\xi_i, \theta_i)$, and the sensing distance (Fig. 3). Line-of-sight, viewshed, visibility, and position of obstacles are essential information, which can be directly obtained from 3D models provided by a geographic information system.

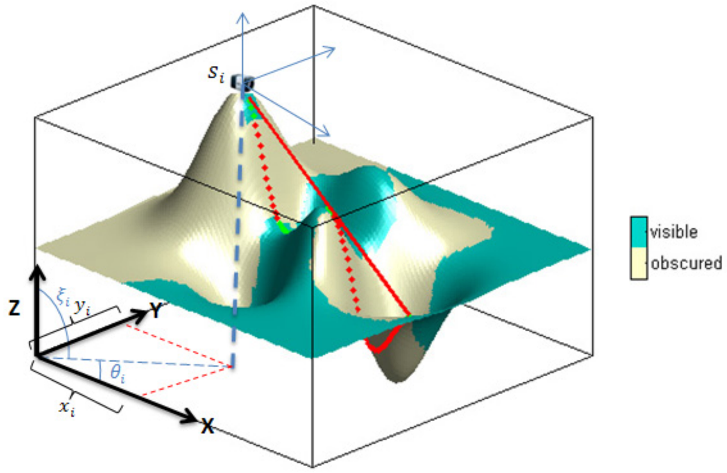



**Fig. 3.** Parameters and visibility of sensor $s_i$ in a 3D model

To find the optima in the category of global optimization methods, it is necessary to define a coverage function. This function is expressed based on the properties of the sensors and the environment information. The sensing model in our investigation is related to distance between sensor and target locations, sensor orientation, and visibility. If we assume that $p_i = (x_i, y_i, z_i)$ is a sensor position, $\theta$ is pan angle around its vertical axe and $\xi$ is the tilt angle around the horizontal axes then, the coverage function $C(s_i, q)$ for sensor $s_i$ at point $\boldsymbol{q}$ can be defined as a function of distance $d(s_i, q) = \|p_i - q\|$, pan angle $p(s_i, q) = \angle_p(q - p_i) - \theta_i$, tilt angle $t(s_i, q) = \angle_t(q - p_i) - \xi_i$, and visibility $v(s_i, q)$ from the sensor (Akbarzadeh et al. 2011):

$$C(s_i, q) = f\left[\mu_d(\|p_i - q\|), \mu_p\left(\angle_p(q - p_i) - \theta_i\right), \mu_t(\angle_t(q - p_i) - \xi_i), v(p_i, q)\right] \qquad (1)$$

Where $\angle_p(q - p_i) = \arctan\left(\frac{y_q - y_{p_i}}{x_q - x_{p_i}}\right)$ is the angle between the sensor $s_i$ and the point $\boldsymbol{q}$ in the horizontal plane and $\angle_t(q - p_i) = \arctan\left(\frac{z_q - z_{p_i}}{\|p_i - q\|}\right)$ is the angle between the sensor $s_i$ and the point $\boldsymbol{q}$ in the vertical plane. Parameters $\mu_d, \mu_p, \mu_t \in [0,1]$ represent membership functions that need to be defined according to the coverage conditions.

In order to cover point $\boldsymbol{q}$ by sensor $s_i$, we should consider the sensing range, sensing angle and visibility. These three parameters can be extracted directly from the 3D model which will be used to make the optimization. The parameters $d(s_i, q)$ and sensing range are calculated based on the $(x, y, z)$ coordinates of the sensor and target location points which are provided by 3D model. Pan $p(s_i, q)$ and tilt $t(s_i, q)$ angles are characteristics which are related to the orientation of the sensor as well as the distance to point $\boldsymbol{q}$ which is calculated from the 3D model.

As described in section 2.2, the enhanced VOR algorithm is a local optimization method, which attempts to move sensors and "heal" uncovered areas. In each step of iteration, visibility, and then viewshed are calculated based on the line-of-sight between the sensors and targets. The covered area for each sensor corresponds to the intersection of its sensing range and the viewshed area. As mentioned before, visibility and viewshed are obtained from 3D models of the study area. Hence, coverage is also affected by the quality of the 3D models.

# 3- Spatial Data Quality in 3D City Models

The deployment of a sensor network in an urban area requires the use of 3D city models. A 3D city model may contain building models, water bodies, transportation objects, vegetation, and city furniture. The building model is the most detailed and frequently used thematic concept of a city model. Different types of buildings may exist in city models, e.g. residential, public, and industry with different details, height, shapes, and volumes. Usually, transportation objects are represented as a linear network in 2D models, but they are geometrically described by 3D surfaces in 3D urban models. In 3D models, a traffic area accompanying the auxiliary objects and obstacles, which bar or affect the traffic transportation, can depict roads. Vegetation features are important components in 3D models which help us to recognize the surrounding environment. They can be represented as single vegetation objects or plant cover (multi solid) objects in 3D city models. City furniture are movable objects such as traffic lights, signs, flower buckets, benches, and bus stops which can be found in residential, traffic, and public area. Spatial location recognition to install sensors can be improved by taking into account these city furniture details in the 3D city models.

The quality of spatial data in a 3D city model that may be used in sensor networks optimization algorithms can undermine its efficiency. According to the spatial data quality literature, spatial data quality depends on several factors; the "internal" quality of spatial data is determined by its actuality, geometric and semantic accuracy, genealogy, logical consistency, and the completeness of the data. This view reflects the producer's perception of quality, which differs from the notion of "external quality". External quality is focused on "fitness for use"; it is defined as the level of fitness between the data and the needs of users (Mostafavi et al. 2004, Devillers et al. 2006).

There has been a consensus about the criteria of internal quality between the ISO, FGDC, and CEN to use the same criteria for geospatial data quality (Devillers et al. 2006). ISO 19113 (Quality principles) and ISO 19114 (Quality evaluation procedures) are two pairs of standards which define the principles for describing the quality. The ISO 19113 recommends five groups to subdivide the data quality elements which can contain quantitative information. These criteria are completeness, logical consistency, positional accuracy, temporal accuracy, and thematic accuracy (Kresse and Fadaie 2004). The ISO 19115 (Metadata) provides the procedures for quality evaluation by defining a dictionary for the data quality elements. According to ISO 19115, metadata contains both quantitative and non-quantitative information. The ISO 94 (Quality management and quality assurance - Vocabulary) addresses the external quality elements. Investigations on the criteria of external quality have been limited to just a few authors. Among them, Wang and Strong (1996) propose four groups for external quality dimensions: intrinsic data quality, contextual data quality, representation data quality, and accessibility data quality. Bédard and Vallière (1995) have investigated the external quality for geospatial data and mentioned these categories as the quality elements of geospatial dataset: definition, coverage, lineage, precision, legitimacy, and accessibility. Oort (2006) has done a comprehensive study on data quality description and applications. He has defined essential terms of spatial data quality and introduced variable methods of investigating the accuracy and errors in spatial and land cover classification. The studies presented so far have mostly considered 2D models. Although, they can be used for 3D models, Walter (2007) has conducted a research on quality control of 3D geospatial data. He mentioned the spatial data quality elements that have a clear meaning in 3D models. He also proposed an automatic update method for the quality control of 3D models composed of laser data, aerial and terrestrial images. His approach is processed with an image interpretation algorithm in order to control the existing objects and find new objects that are not in the database.

As discussed earlier, it is too difficult to find an exact investigation on the elements of data quality in 3D models. So, in the following list we propose the most relevant criteria of data quality for 3D spatial models.

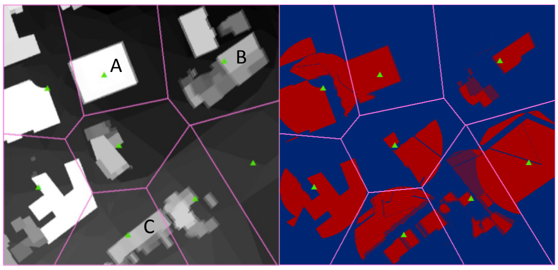
- **Positional Accuracy:** In general, accuracy addresses the probable differences between the measured and true values. It can be divided into relative and absolute accuracy. Positional accuracy is the accuracy of coordinate values and categorized as vertical and horizontal. In 3D city models, compared to 2D models, apart from X and Y coordinates, Z values should be considered in positional accuracy analysis. For example, the accuracy of the height of buildings and other 3D objects is important as well as horizontal positions and it has a direct impact on 3D issues such as shadow and visibility analysis.
- **Logical Consistency:** Logical consistency of a spatial database constitutes an important part of the determination of the internal spatial data quality. It may be defined as the degree of consistency of the data with respect to its specifications. It concerns the question of whether collected data are related to other data in a logical sense. In the other words, it refers to the absence of apparent contradictions in a database (Walter 2007). For instance, in 3D datasets, logical consistency can refer to topological relations. For example, extracting building footprints and extruding them is one of the simplest methods to construct 3D city models from 2D data. So, if the topological relations between the footprints are not taken into account, the resulting 3D city model may not be topologically and hence logically consistent (Ledoux and Meijers 2009).
- **Lineage** concerns the question of how the data are collected and the method of how the data have been entered in a computer program. This information contains a short history of the data producer, data source, data capturing, and data processing methods. In 3D datasets, the lineage can refer to the historical information about data acquisition, data representation and data processing. The question of which kind of instrument or acquisition method has been used to collect the dataset will be answered in the data acquisition part of the lineage information. In data representation we will find the method by which the dataset has been represented, e.g. regular grid, TIN, mesh, 3D faces. In terms of data processing, lineage may contain information about processing methods such as different kinds of interpolation (e.g. nearest neighbour, bilinear, and bicubic). This may also describe the methods used for 3D modeling process.
- **Semantic Accuracy** addresses the question of whether the data really express its intended meanings. This criterion provides information on the difference between the values of spatial attributes and their real values. In 3D models, we are again concerned with the semantics of spatial objects. More specifically we are concerned with the semantics of 3D objects. For example how to represent buildings regarding their definitions and shapes (factory, hospital, residential, educational, etc.). What are the spatial integrity constraints that exist between 3D objects and how accurately are they defined with respect to the reality?
- **Completeness** indicates the question of whether there is anything more to add to the data. This criterion is usually determined based on the matrix of omission (abnormal absence) and commission (abnormal presence) of some objects in a spatial model. This can also be related to the levels of detail (LOD) used to represent spatial information in 3D models. Omission or commission of some objects (e.g. trees and buildings) or the levels of detail in their representation (e.g. missing balconies in a 3D building model) in a 3D city model have an impact on the sensor positions obtained from the optimization algorithms. Also, it will have a significant impact on the estimation of its spatial coverage.

- **Temporal Accuracy** concerns the question of whether the data is up to date or not. For example are there some new constructions in a 3D city model, which should be added to the dataset, or it is necessary to delete some blocks from the dataset.

## *4- The Impact of 3D City Models Quality on Sensor Deployment*

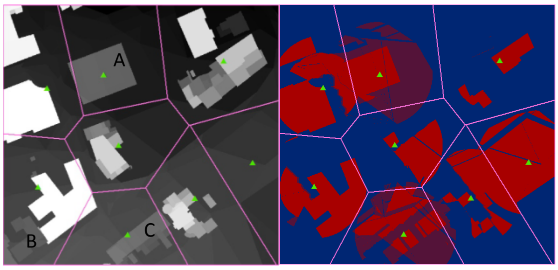
As mentioned in Section 2, the sensor placement optimization algorithms that are applied in our experimentation use line-of-sight and viewshed to calculate spatial coverage. These two concepts allow visible and invisible objects to be identified and hence, define covered and uncovered areas in the region of interest. The quality of 3D city models has a direct impact on the estimation of these values. In the following, we will present and discuss these impacts with respect to some of the quality criteria that we presented in the previous section.

Positional accuracy has a direct impact on the estimation of the visibility in a 3D city model. The positional accuracy may be presented as a small displacement in the position of the objects, which can be either horizontal or vertical or both. Even a few centimeters inaccuracy in horizontal or vertical positions of objects or sensors can block the line-of-sight between a sensor and a target. Fig.4 shows the impact of changing the positions of buildings on the obtained coverage. In Fig. 4(a) positions of three buildings have been displaced and overlaid at the same DTM. So, buildings at points A, B, and C are opaque. Fig. 4(b) depicts the change of coverage because of inaccuracy in the positions of those buildings. Fig. 5 shows the impact of completeness on sensor network coverage. In Fig. 5(a), three buildings have been removed from the dataset. Fig. 5(b) depicts the impact of elimination on the final coverage. This situation may also occur in datasets due to temporal accuracy and the demolishing of some buildings. Conditions have been considered in an exaggerated manner in both Fig. 4 and 5.




(a) (b)

**Fig. 4.** Impact of positional accuracy on sensor network coverage: (a) small displacement of three buildings at positions A, B, and C shown on DTM (b) area which will be covered after the displacement shown in light red.




**(a)** **(b)**

**Fig. 5.** Impact of completeness on sensor network coverage: (a) elimination of some blocks at positions A, B, and C shown on DTM (b) area which will be covered after elimination shown in light red

In 3D city models, sensor nodes and 3D objects should be logically consistent. In order to respect consistency in sensor network deployment, the topological relationship must be observed. To ensure logical consistency in a model, some logical rules should be defined and then the validity of those rules must be verified in the model. For example a sensor node for monitoring the traffic in a city could not be placed on the top private property or area with a height lower than a predefined threshold. Another example could be poles in the city, which are supposed to be used for installing sensors. If they stand at the right side of the street but they are represented at the left side in 3D city model, our model is not consistent with the reality and optimization with this model would have a significant impact on the coverage. The maximum distance of coverage and communication between sensors should also respect logical rules when we try to place sensors in the environment.

In addition to the importance of accurate geometrical and topological representations of spatial information in 3D city models, semantic accuracy of spatial features is also essential for efficient optimization of a sensor network. Semantic accuracy deals with precise definitions of spatial, temporal, and thematic properties of each feature represented in 3D city models. Spatial features such as buildings, streets, poles, transportation objects, water bodies, and vegetation area must be accurately identified, classified and specified in the models. Thematic information must be semantically rich enough to allow consideration of all possible restrictions in the optimization process of sensor deployment in a given urban area.

Another important spatial data quality criterion that may have a significant impact on the coverage estimation of a sensor network is the completeness of spatial information in a 3D city model. As mentioned in the previous section, completeness of data may have different implications in a 3D city model including omissions, commissions and levels of detail (LOD) in the representation of an urban area. Open Geospatial Consortium (OGC) has adopted City Geography Markup Language (CityGML) as a standard for representation of 3D city models (OGC 2008). CityGML introduces five levels of detail to support multi-scale modeling of an urban area. In a 3D city model, the same object may be represented in different LOD simultaneously, enabling the analysis and visualisation of the same object with regard to different degrees of resolution. Hence, spatial representation of an object may have some details in one level that can disappear in another level of detail. The roughest level, LOD0, is a two and a half dimension DTM and may be used for regional and landscape applications. LOD1 is the blocks model in the city or region, which represents buildings with flat roofs. In LOD2, roof structures in buildings are differentiated and vegetation objects may also be shown. City districts may be represented in LOD2. LOD3 contains architectural elements of buildings with detailed walls, roofs and balconies.

Other urban structures such as detailed vegetation and transportation objects may appear in LOD3. LOD4 is a higher resolution representation of LOD3 with information on interior structures of 3D objects. Figure 6 depicts the five levels of detail in an urban area.

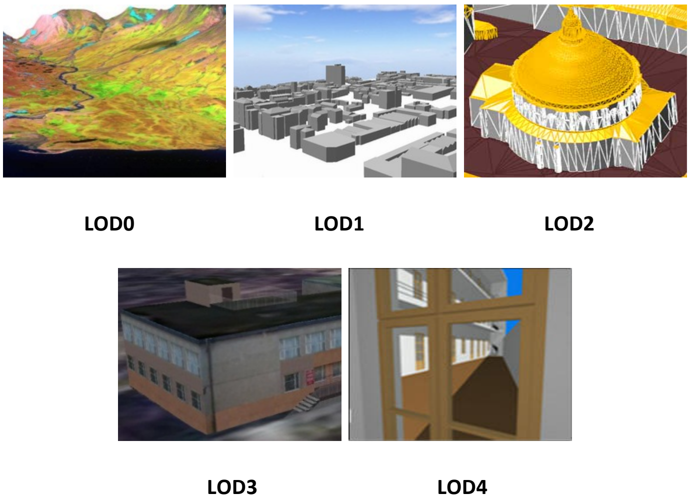



**Fig. 6.** The five Levels of Detail in 3D city models (LOD) according to the Open Geospatial Consortium (Gröger et al. 2012)

In a sensor network deployment, the presence of some details directly affects the line-of-sight measurements and makes a specific target visible or obscured. For example, consider a building with balconies which has been represented in LOD1. Consider also that there is a sensor placed on the top of such a building. As shown in Fig. 7, the omission of the balconies in the 3D representation of the building will result in a complete coverage area compared to the case where the balconies are present in the 3D representation of the building.


**Fig. 7.** The impact of completeness on the line-of-sight; Point B is visible but point A is invisible because of the presence of balconies in the building representation

Another important issue that has a significant impact on the sensor network optimization is the type of spatial representation of the real world. Vector and raster models are two fundamental representation methods of the reality. Vector representations of the reality are often more accurate for spatial features with well-defined limits such as in buildings and streets (Fig. 8). However, most of the optimization algorithms are conceived based on raster representation of the environment since using raster models is less complex than vector models. In addition to the accuracy of representation of 3D objects, sensors could be more accurately positioned in vector maps. Indeed, an accurate determination of sensor positions in a raster representation of the space is more difficult. In addition, we need a very high resolution for 2D or 3D representation of the space in order to achieve the required precision. We think vector representation of the space such as in 3D city models will help to more precisely estimate spatial coverage of a sensor network, because visibility could be estimated more accurately in vector data. However, in our knowledge development of optimization algorithms for sensor networks using vector data models are poorly investigated, and more research work is needed in the field.

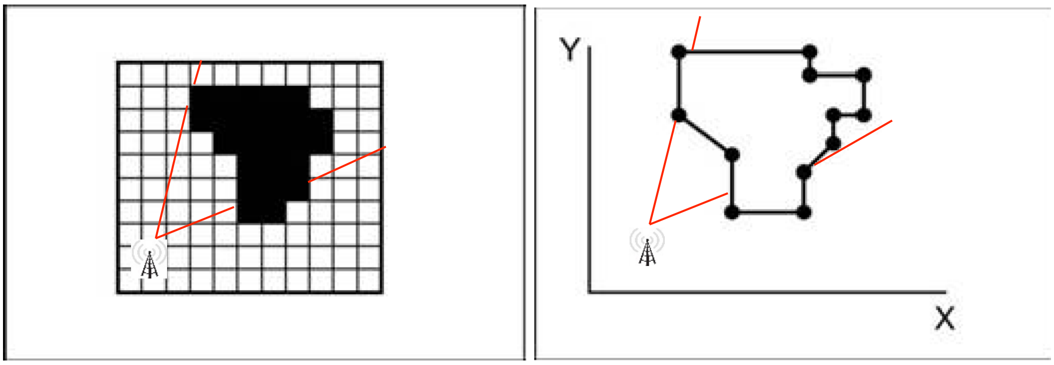



**Fig. 8.** Raster versus vector representation of limits of a building; visibility and line-of-sigh can be computed more accurately using vector model.

## 5- Experimentation and Results

As discussed in the previous section, the quality of 3D city models can be evaluated based on different criteria. In this section, we will carry out different experiments in order to show the impact of the map resolution and sensor configuration of a 3D model in the estimation of spatial coverage of a sensor network. For this purpose, we have prepared 5 maps with different resolutions from the same area. Our goal is to investigate the impact of the positional accuracy and completeness of the dataset on the spatial coverage of a sensor network that will be introduced to the optimization algorithms. Here, the completeness implies the presence of some details in higher resolution maps that are omitted in the maps with lower resolution as discussed in previous sections. The resolution variation is from 500 cm (low resolution) to 50 cm (high resolution) and a map with 10 cm resolution is considered as the ground truth dataset to validate the results. The maps dimension is 180 m by 170 m from an urban area in old Quebec City, Canada. Fig. 9 depicts the 3D model of the study area.

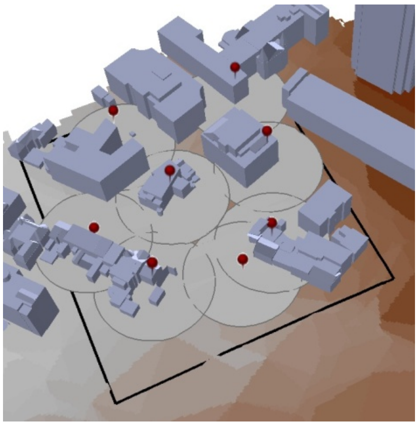


**Fig. 9.** Sensor locations in a 3D city model, red points show the assumptive positions of 8 sensors in the environment and the circles depict their sensing area

This experiment consists of deploying eight sensors inside the study area, in order to obtain the best possible coverage by means of an optimization algorithm. It has been supposed that each sensor has a 35 meters sensing range, positioned one meter height above the surface and has the ability to rotate 360 degrees horizontally and ±90 degrees vertically at its position.

As discussed in Section 2, two different types of optimization algorithms have been used for sensor network deployment: global and local approaches. In order to compare the sensitivity of the proposed optimization algorithms to the input dataset quality, we have chosen three optimization methods. For global approaches we have selected Covariance Matrix Adaptation Evolutionary Strategy (CMA-ES) and Simulated Annealing (SA). Among different stochastic optimization methods, these two methods were chosen because one (i.e. SA) is an example of simple stochastic methods and the other one (i.e. CMA-ES) is an example of more sophisticated methods. Among local methods, enhanced VORonoi-based algorithm has been chosen in order to consider the geometrical characteristics of the study area. For all methods, each sensor placement optimization scheme was run 32 times, from which the average of each method was estimated. The initial positions of the sensors were determined randomly for each method.

To assess the sensitivity of the optimization methods with respect to the quality of data set, we have conducted several experiments and presented the results in Tables II, III and IV. The experiments were carried out as follow:

1- First, for the purpose of the experiments, five maps with different resolutions from the same area were created;
2- For each map, we have conducted the optimization process using three methods as mentioned above;
3- Then, we have computed the average and best coverage values for each map using each of those optimization methods (columns 2 and 3 in Tables II, III, and IV);
4- Next, the best sensor configuration is selected for each map based on best coverage;
5- Finally, for each best sensor configuration, the best and average coverage values were computed on the ground truth (columns 4 and 5 in Tables II, III, and IV).

It should be mentioned that in order to be able to compare the obtained results from the experiments, we applied the same sensing range for all sensors. The sensing area for each sensor was considered to be a crisp circle. In addition, the same algorithms were used for the determination of visible and non-visible pixels and the coverage values for all the optimization methods. The performance of each method has also been evaluated by defining the same function for computing the viewshed inside the study area. Table I presents the configurations which have been used for the experiments.

**Table I:** Initial information on the sensor network used in our experimentation.

| *Method* | *Num. of Sensors* | *Sensing range (m)* | *Num. of runs* | *Max. iteration* |
|---|---|---|---|---|
| CMA-ES | 8 | 35 | 32 | 300 |
| SA | 8 | 35 | 32 | 4200 |
| VOR | 8 | 35 | 32 | 200 |

The results for CMA-ES have been reported in Table II, SA in Table III and, enhanced VORonoi-based in Table IV.

**Table II:** Results obtained from the CMA-ES method.

| *Resolution (cm)* | *Avg. coverage (%)* | *Best coverage (%)* | *Best coverage over 10cm resolution (%)* | *Average coverage over 10cm resolution (%)* |
|---|---|---|---|---|
| 500 | 52.50 | 52.96 | 44.79 | 45.09 |
| 300 | 52.78 | 53.79 | 46.62 | 47.75 |
| 200 | 49.09 | 51.33 | 43.85 | 46.34 |
| 100 | 50.75 | 52.77 | 41.27 | 46.50 |
| 50 | 50.75 | 52.72 | 52.50 | 47.85 |

**Table III:** Results obtained from the SA method.

| *Resolution (cm)* | *Avg. coverage (%)* | *Best coverage (%)* | *Best coverage over 10cm resolution (%)* | *Average coverage over 10cm resolution (%)* |
|---|---|---|---|---|
| 500 | 45.50 | 51.73 | 47.40 | 40.06 |
| 300 | 45.16 | 49.98 | 46.10 | 40.97 |
| 200 | 42.59 | 48.95 | 49.09 | 41.28 |
| 100 | 45.75 | 48.07 | 42.33 | 41.85 |
| 50 | 44.97 | 47.55 | 47.12 | 43.35 |

**Table IV:** Results obtained from the enhanced VORonoi-based method.

| *Resolution (cm)* | *Avg. coverage (%)* | *Best coverage (%)* | *Best coverage over 10cm resolution (%)* | *Average coverage over 10cm resolution (%)* |
|---|---|---|---|---|
| 500 | 45.55 | 47.19 | 43.14 | 42.21 |
| 300 | 47.83 | 51.07 | 45.87 | 45.37 |
| 200 | 40.06 | 43.82 | 42.43 | 40.51 |
| 100 | 44.38 | 45.77 | 44.25 | 42.83 |
| 50 | 46.59 | 48.16 | 45.64 | 44.32 |

The goal of the comparison between the three algorithms is not to determine which algorithm outperforms the other methods; our objective in this investigation is to discover the sensitivity of optimization algorithms to the quality of input datasets. Results presented in Fig. 10 show that all three methods have good stability regarding the inaccuracy of the input dataset (between 5 meters to 50 centimeters resolution). CMA-ES gives better coverage in all resolutions while SA and VOR have returned almost the same results. We presume that the reason is that CMA-ES is a more sophisticated optimization method, which derives a second order model of the objective function while SA is a simple stochastic optimization method, which randomly searches for a better solution in the search space. All three methods suffer the worst results when the resolution is 200 cm. The reason is related to the shape of objects in this study area. Comparing other resolutions, building details and real shapes begin to appear in 200 cm resolution, which causes more area to be obscured from sensor visibility. At lower resolutions, optimization algorithms perform better due to the disappearing of obstacles in the datasets. For higher resolutions, the scenario is changed; optimization algorithms perform better due to their inherent process to find the optimum when the pixel size is smaller. So, in this study area the resolution of 200 cm could be considered as a critical resolution.

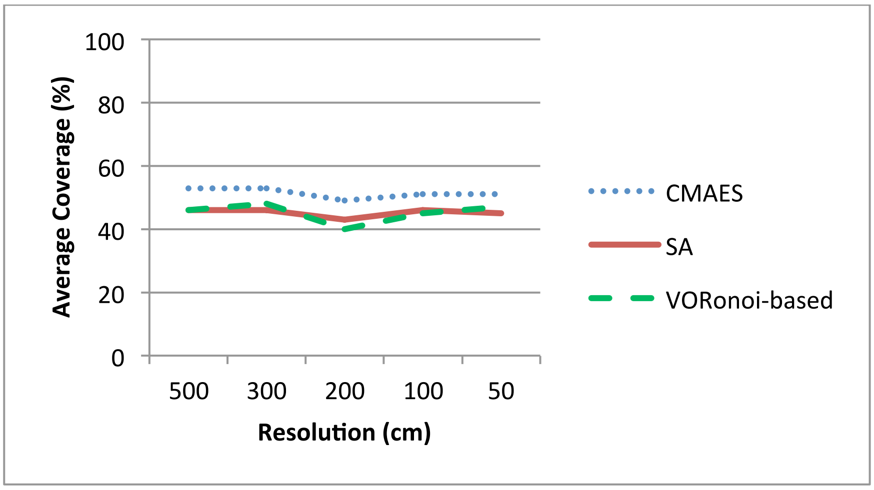



**Fig. 10.** Comparison of the accuracy of three optimization methods tested regarding the accuracy of the input datasets

Fig. 11 compares the configuration of sensor positions and related coverage obtained over a map with 10cm resolution by using three optimization methods. The sensor positions that are obtained from CMA-ES give 52.50% coverage over the study area, while sensor configurations obtained from SA and VOR methods give 47.12% and 45.64% coverage. The sensors have been positioned in almost the same places in all three algorithms with a few differences, which means all algorithms have located almost the same places to place sensors with different input data quality.

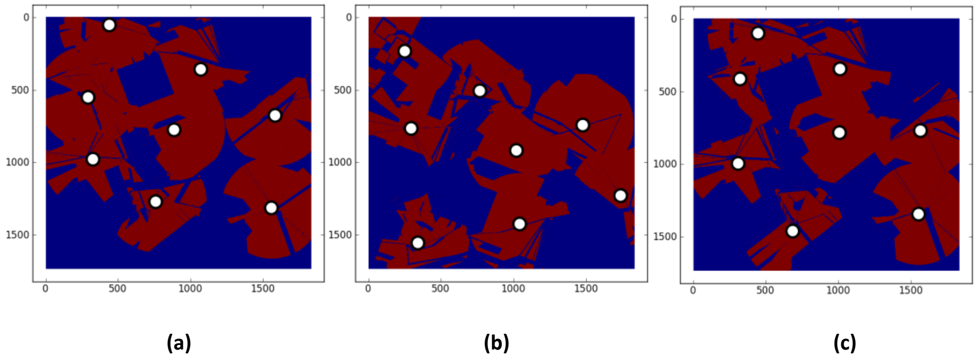



**Fig. 11.** (a) Best sensor configuration over a map with 10 cm resolution from CMA-ES, (b) best sensor configuration over a map with 10 cm resolution from SA, and (c) best sensor configuration over a map with 10 cm resolution from enhanced VORonoi-based.

Fig. 12 depicts the sensitivity of optimization methods with respect to the accuracy of input data. To do so, we calculated the differences between the averages of coverage for each map and the coverage evaluated by applying the best configuration of the sensor positions of all runs obtained from each map over the ground truth data.

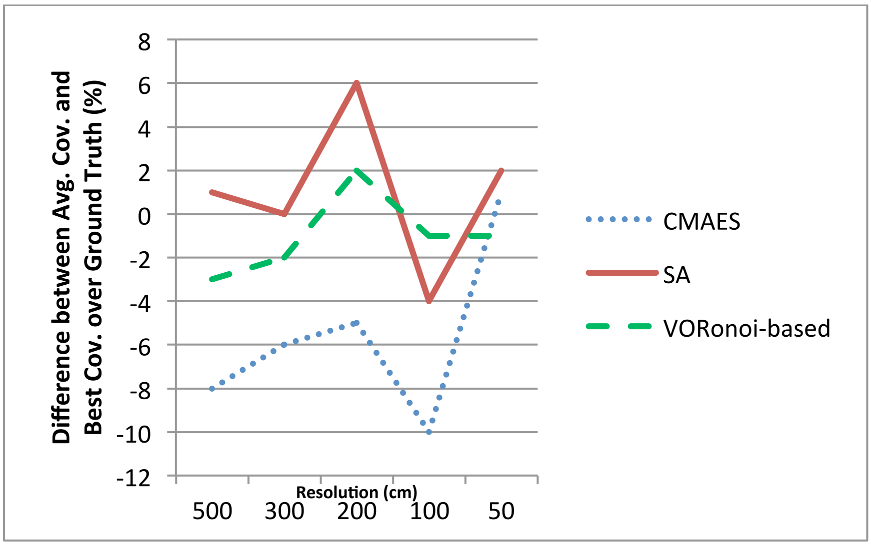



**Fig. 12.** Comparison of differences between average coverage obtained from different optimization methods over different map resolutions and best coverage obtained from different algorithms and map resolutions over ground truth dataset.

As illustrated in Fig. 12, best coverage evaluated over 10 cm resolution map by using the best sensor configuration results has been influenced by the input data quality (resolution). To investigate this impact more accurately, we have compared the average of coverage for each map and the average of coverage evaluated over the ground truth data for all runs. Fig. 13 shows sensitivity of different optimization algorithms with respect to the accuracy of input datasets. We can see from the figure that as the resolution of the maps becomes higher the difference between the sensitivity of the optimization algorithms becomes smaller. We also observe that there is a peak in all curves at 200 cm resolution for all the optimization methods. As discussed earlier, this behaviour is related to the worst coverage in that resolution, which does not exist in the evaluated coverage of ground truth data.

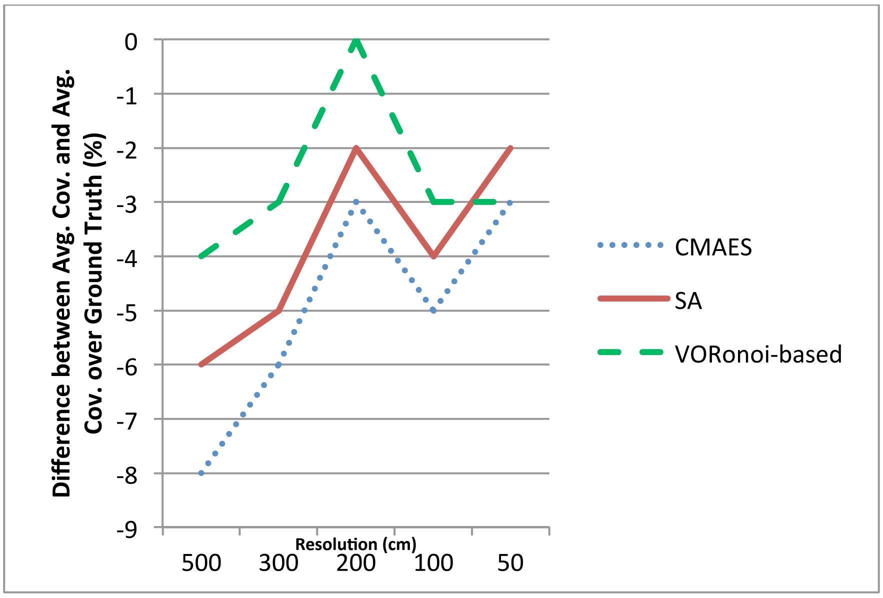



**Fig. 13.** Comparison of differences between average coverage values obtained from original maps and the ground truth for three optimization methods.

Table V and Fig. 14 show the results of the one-way analysis of variance (ANOVA) for three methods over maps with different resolutions. The one-way ANOVA is a technique to compare the means of samples to test whether those samples in two or more groups are drawn from the same population or not. The ANOVA produces a F-statistic, the ratio of the variance calculated among the means to the variance within the samples. If the group means are drawn from the same population, the variance between the group means should be lower than the variance of the samples. A higher ratio, therefore, implies that the samples were drawn from different populations (Hogg and Ledolter 1987). In this study, we used ANOVA to determine to what extent our evaluations of coverage over maps with different resolutions differ from each other. If we assume that all optimization methods should report the same coverage by using different resolution maps from the same study area, the F-statistic value allows us to determine to what extent coverage values are similar. In this case study, given that we have 5 groups (5 maps with different resolutions) and 32 samples per groups (32 runs for each map), the maximum F-statistics $F(x,y)$, with $x=4$ (5-1) and $y=128$ ($32\times(5-1)$) for a probability level of 0.05 which allows us to test whether the results have a 95% chance of coming from the same statistical population, would be $F(4,128)=2.44$. So, a greater F-statistic refers to a higher sensitivity to the quality of the input dataset and vice versa. Table V indicates that the differences in the F-statistics results obtained by varying the resolution is significant for all methods, which confirms the sensitivity of all methods to the quality of input datasets. The lowest F-statistic 5.44 in Table V was obtained for SA, which indicates that the average coverage values from SA have more likely been obtained from the same populations, and therefore, SA, is less sensitive. The highest F-statistic is 229.6 for VOR, which is thus more sensitive to the quality of input data. The reason is that SA uses small absolute displacement to determine the optimum positions, which is not related to the resolution. The box plots of the different map resolutions for each method in Fig. 14 indicate that the standard deviations with VOR are lower when compared to CMA-ES and SA, which indicates that results obtained from VOR algorithm are coherent in each run for the same map resolution. The reason is that VOR is a deterministic algorithm that uses the geometric structure of the environment, which is not changed by changing the initial sensor positions. So, applying the algorithm with different initial starting positions for sensors has less impact on the final results. Conversely, SA is a highly stochastic algorithm, which returns the highest standard deviation in the results for different runs on the same map resolution. CMA-ES is in-between these two algorithms, being more stable than SA, but still gives results with a higher standard deviation than VOR since it is also a stochastic optimization algorithm.

**Table V:** F-statistic results from one-way ANOVA test.

| | *ANOVA F-statistic* |
|---|---|
| CMA-ES | 23.78 |
| SA | 5.44 |
| VOR | 229.6 |

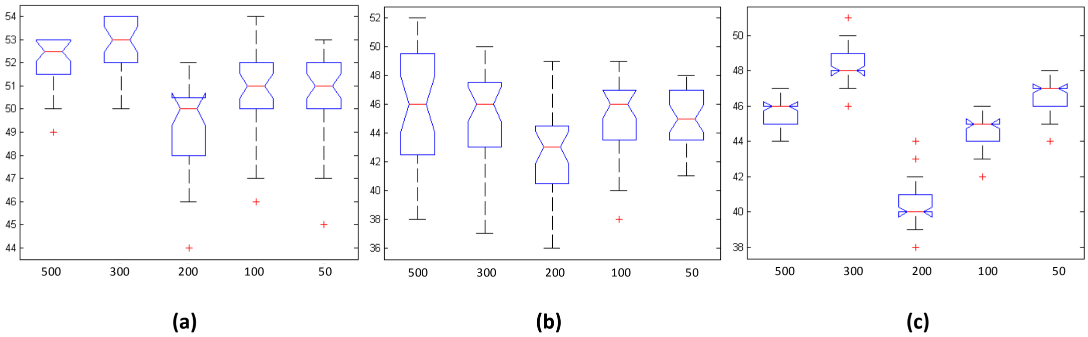


**Fig. 14.** Box plot of one-way ANOVA test for different map resolutions for: (a) CMA-ES; (b) SA; (c) VOR methods

## *6- Conclusions*

A survey on spatial data quality in 3D city models was conducted in this paper and a list of the most relevant elements of data quality for 3D models was proposed. The impact of 3D data quality elements on sensor placement has been determined by investigating their impact on the concepts of viewshed and line-of-sight. Positional accuracy and completeness were introduced as two important elements in sensor network deployment. The concepts of raster and vector data and their accuracies when are used as input in sensor network optimization algorithms were discussed. To examine the impact of 3D data quality on sensor network placement and calculated coverage, a comparison of the sensitivity between three optimization algorithms on the quality of input data was carried out. The algorithms which were used in this investigation were some global and local optimization methods with the novelty of integration of 3D models. The impact of data quality on final coverage and sensitivity of each method was studied by using different maps with different quality as input data to the optimization algorithms. Map resolutions range from 500 to 50 cm and a map of 10 cm resolution considered as the ground truth data.

The results show that all methods are slightly stable with different resolution, which indicates that both global and local optimization algorithms are less sensitive to the quality of input data and return almost the same results. Regarding the algorithm of SA, it is less sensitive when compared to others, however the deviation is higher in the final coverage results. VOR has less deviation but it is a little more sensitive to the quality of input data. In terms of final coverage, CMA-ES performed better than the SA and enhanced VOR algorithms.

This research is not exhaustive in terms of studying the sensitivity of optimization algorithms with respect to all the data quality criteria. The research is however significant in terms of proposing a methodology for the assessment of the sensitivity of an optimization method with respect to the quality of spatial 3D models (ex. 3D city models). Throughout the paper, we have defined and illustrated the impact of some of the 3D data quality elements on the estimation of sensor network coverage. And finally, we have carried out some experimentation three most reliable and used optimization algorithms to illustrate more concretely the impact of the quality of 3D city models on the estimation of coverage in urban areas. Further investigations are required to define and analyse the impact of the spatial data quality for each quality criterion on the estimation of the spatial coverage of a given sensor network. It would be also interesting to carry out new experimentation on the quality assessment of 3D datasets with higher LODs for an urban area.

## *Acknowledgments*

This research was supported by the project entitled *Integrating developmental genetic programming and terrain analysis techniques in GIS-based sensor placement systems* funded by the Canadian GEOIDE Network of Centres of Excellence, MDA Systems Ltd and DRDC Valcartier. We also thank Annette Schwerdtfeger for the proofreading of this paper.

## *References*

Ahmed N., Kanhere S. S., Jha S. (2005) The holes problem in wireless sensor networks: A survey. ACM SIGMOBILE Review, vol. 9, no. 2.

Akbarzadeh V., Gagné C., Parizeau M., Mostafavi M. A. (2011) Black-box optimization of sensor placement with elevation maps and probabilistic sensing models, in Proc. Of the IEEE International symposium on robotic and sensors environment (ROSE 2011)

Akbarzadeh V., Ko A., Gagné C., Parizeau M. (2010) Topography-Aware Sensor Deployment Optimization with CMA-ES, in Proc. of Parallel Problem Solving from Nature (PPSN 2010), Springer

Argany M., Mostafavi M. A., Karimipour F. (2010) Voronoi-based Approaches for Geosensor Networks Coverage Determination and Optimisation: A Survey, in Proc. ISVD 2010, Quebec City

Argany M., Mostafavi M. A., Karimipour F., Gagné C. (2011) A GIS Based Wireless Sensor Network Coverage Estimation and Optimization: A Voronoi Aproach. Journal of Transaction on Computational Science XIV, LNCS 6970, 151-172, Springer-Verlag Berlin Heidelberg

Aziz N. A., Aziz K., Ismail W. Z. W. (2009) Coverage strategies for wireless sensor networks, World Academy of Science, Engineering and Technology, vol. 50, 145-150.

Bédard Y., Valliére D. (1995) Qualité des données à référence spatiale dans un contexte gouvernemental, technical report, Quebec City, Laval University.

Bharathidasan A. and S. Madria (2003) Sensor networks: an overview, IEEE Potentials, vol. 22, pp. 20–23.

Devillers R., Jeansoulin R. (2006) Spatial data quality: concepts, book chapter, Fudumentals of Spatial Data Quality, ISTE, 31-42

Ghosh A., Das S. K. (2008) Coverage and connectivity issues in wireless sensor networks: a survey, Pervasive and Mobile Computing, vol. 4, 303–33.

Gröger G., Kolbe T. H., Nagel C., Häfele K. H. (2012) OGC City Geography Markup Language (CityGML) Encoding Standard, OGC, USA

Hansen N. and Ostermeier A. (2001) Completely derandomized self-adaptation in evolution strategies. Evolutionary Computation, vol. 9(2), 159-195

Hogg R. V., and Ledolter J. (1987) Engineering Statistics. New York: MacMillan,

Huang C. F. and Tseng Y. C. (2005) A survey of solutions to the coverage problems in wireless sensor networks, J. Internet Tech., vol. 6, no. 1, 1-8.

Kirkpatrick S., Gelatt C. D., and Vecchi M. P. (1983) Optimization by simulated annealing. Science, 220(4598):671

Kresse W. and Fadaie K. (2004) ISO Standards for Geographic Information (first ed), Springer, Heidelberg (322pp.)

Ledoux H. and Meijers M. (2009) Extruding building footprints to create topologically consistent 3D city models. In A. Krek, M. Rumor, S. Zlatanova, and E. Fendel, editors, Urban and Regional Data Management, UDMS Annual, CRC Press, 39-48.

Niewiadomska-szynkiewicz E., Marks M. (2009) Optimization schemes forwireless sensor network localization, Int. J. Appl. Math. Comput. Sci., vol. 19, no. 2, 291–302

Nittel S. (2009) A survey of geosensor networks: advances in dynamic environmental monitoring, Sensors Journal, vol. 9, 5664-5678.

Megerian S., Koushanfar F., Potkonjak M., and Srivastava M. B. (2005) Worst and best-case coverage in sensor networks, IEEE Transaction on mobile computing, vol. 4, no. 1

Meguerdichian S., Koushanfar F., Qu G., and Potkonjak M. (2001) “Exposure in wireless ad-hoc sensor network,” Proc. IEEE MOBICOM ’01, 139-150.

Mostafavi M. A., Edwards G., Jeansoulin R. (2004) Ontology-based method for quality assessment of spatial data bases. in: Proceedings of ISSDQ'04, GeoInfo Series, Bruck/Leitha, Austria, 49-66.

Oort V. P. (2006) Spatial data quality: from description to application, PhD Thesis, Wageningen: Wageningen Universiteit

Romoozi M., Ebrahimpour-komleh H. (2010) A positioning method in wireless sensor networks using genetic algorithms, International Journal of Digital Content Technology and its Applications vol. 4, 174-179.

Walter V. (2007) Quality Control of 3D Geospatial Data, Photogrammetric Week ’07, Stuttgart University, 315-324

Wang R. Y. and Strong D. M. (1996) Beyond Accuracy: What data quality means to data consumers. Journal of Management Information System, vol. 12, Issue 4, 5-33

Wang Y. C. and Tseng Y. C. (2008) Distributed deployment schemes for mobile wireless sensor networks to ensure multilevel coverage. IEEE Transactions on Parallel and Distributed System, vol. 19, Issue 9, 1280-1294