

ACCURACY EVALUATION OF A MOBILE MAPPING SYSTEM WITH ADVANCED STATISTICAL METHODS

I. Toschi ^a, P. Rodríguez-Gonzálvez ^b, F. Remondino ^a, S. Minto ^a, S. Orlandini ^c, A. Fuller ^d

^a 3D Optical Metrology (3DOM) unit, Bruno Kessler Foundation (FBK), Trento, Italy - (toschi, remondino, minto)@fbk.eu

^b Dept. of Cartographic and Land Engineering, High School of Ávila, University of Salamanca, Ávila, Spain - pablorgsf@usal.es

^c Microgeo, Firenze, Italy - simone@microgeo.it

^d RIEGL, Vienna, Austria

Commission V, WG V/4

KEY WORDS: Mobile, Accuracy, Precision, Statistic, Terrestrial, Photogrammetry, Laser scanning

ABSTRACT:

This paper discusses a methodology to evaluate the precision and the accuracy of a commercial Mobile Mapping System (MMS) with advanced statistical methods. So far, the metric potentialities of this emerging mapping technology have been studied in few papers, where generally the assumption that errors follow a normal distribution is made. In fact, this hypothesis should be carefully verified in advance, in order to test how well the Gaussian classic statistics can adapt to datasets that are usually affected by asymmetrical gross errors. The workflow adopted in this study relies on a Gaussian assessment, followed by an outlier filtering process. Finally, non-parametric statistical models are applied, in order to achieve a robust estimation of the error dispersion. Among the different MMSs available on the market, the latest solution provided by RIEGL is here tested, i.e. the VMX-450 Mobile Laser Scanning System. The test-area is the historic city centre of Trento (Italy), selected in order to assess the system performance in dealing with a challenging and historic urban scenario. Reference measures are derived from photogrammetric and Terrestrial Laser Scanning (TLS) surveys. All datasets show a large lack of symmetry that leads to the conclusion that the standard normal parameters are not adequate to assess this type of data. The use of non-normal statistics gives thus a more appropriate description of the data and yields results that meet the quoted *a-priori* errors.

1. INTRODUCTION

1.1 MMS technology

Mobile Mapping System (MMS) is nowadays an emerging technology, whose development began in the late 1980s and is constantly growing. From a technological point of view, MMS is a multi-sensor system, that consists mainly of three components: mapping sensors (active and/or passive 3D imaging systems), navigation/positioning sensors (IMU/GNSS) and a control unit, that synchronizes and integrates the acquisition of geometric/positioning information. All sensors are integrated on a rigid moving platform (e.g. vans, cars, trains, boats, snow mobile sledges, people, etc.), whose trajectory is computed and finally used to produce geo-referenced 2D/3D data. Land-based mobile laser scanners mounted on vans or cars represent the best and a cost-effective solution for capturing 3D point clouds of urban areas with an high recording rate, high point density/accuracy and remote acquisition mode. Thanks to these benefits and to continuous developments in both scanning and navigation technologies, MMSs are gaining more and more importance in many application fields, such as civil engineering and construction (Slattery et al., 2012), environment (Bitenc et al., 2011), pipeline design (Kawashima et al., 2012), road inventory (Pu et al., 2011) and cultural heritage (Ziparo et al., 2013). A general review of the different solutions available on the market and a comparison of their technical specifications can be found in Puente et al. (2013).

1.2 Literature review on approaches for accuracy evaluation of MMS

Terrestrial mobile mapping technology goes back to the 90's when the first experiments showed the potential of mobile acquisitions for mapping purposes (Bossler and Novak, 1993; El-Sheimy et al., 1995; El-Sheimy, 1996).

MMSs have then developed from a typical academic research to commercially operating systems used for topographic surveying, 3D mapping of traffic arteries, city planning, visual street-level image and vector data acquisition, visualization, etc. The accuracy requirements for the acquired data is substantially different in each application, with cartographic mapping and road or rail infrastructure surveying being much more demanding in this particular respect. Nowadays there are different commercial MMS solutions and they show the best example of sensor integration and cost-effective acquisition of geo-referenced spatial data, with a combination of digital imaging devices, long-range laser scanners and GNSS/IMU positioning sensors (Tao and Li, 2007). Besides results published by system manufacturers and linked to internally-performed tests/procedures, there are few publications where MMS performances are examined with an external reference dataset and by an independent actor.

One of the most recent studies has been carried out within the European Spatial Data Research (EuroSDR) project "Mobile Mapping – Road Environment Mapping Using Mobile Laser Scanning" (Kaartinen et al., 2012). This action aimed at benchmarking the performance of various MMSs using a permanent urban test-field established around a shopping mall. A digital elevation model, some poles and building corners were adopted as reference objects. After comparing the collected datasets against the ground-truth and performing a 'gross error filtering', the accuracy was finally derived by computing the minimum, the maximum and the standard deviation values of the differences. Tests provided an elevation accuracy better than 3.5 cm for all professional systems properly calibrated (up to a range of 35 m), whereas the best system achieved a planimetric accuracy of 2.5 cm up to a range of 45 m.

Lim et al. (2013) performed a field test at Padre Island National Seashore (USA) in order to evaluate the accuracy of point clouds collected by a MMS under variable settings of vehicle speed and point density. Six vertical targets were GNSS-

measured and adopted as references. Comparisons led to a mean target error of -0.222 m (East), -0.036 m (North) and -0.104 m (Height). Given the excellent conditions for GNSS provided by the test-area, discrepancies were ascribed to boresight errors and a parameter-domain approach was developed in order to adjust the lever arm parameters.

Prior to these studies, the StreetMapper mobile system received different attentions. Kremer and Hunter (2007) declared an accuracy range of 3-5 cm and reported some practical applications where the MMS could be used for mapping purposes. Barber et al. (2008) used GNSS measurements as references to evaluate precision and accuracy of laser scanning data collected in two test-sites, i.e. a residential and an industrial area. RMS errors in elevation were found to be in the order of 3 cm, whereas the planimetric accuracy was set around 10 cm. Haala et al. (2008) adopted an existing 3D city model of the city of Stuttgart to provide area covering reference surfaces. The investigation sought to assess the quality of data collected at building façades. After a semi-automatic selection of surfaces from the 3D city model, corresponding planar patches were estimated by least square adjustment from the MMS-collected data. Comparisons proved that an accuracy better than 3 cm (standard deviation of the differences between measured and reference data) can be achieved by the system in good GNSS conditions.

1.3 Paper objectives

This paper describes an approach designed to analyse the precision and the accuracy achieved by a commercial MMS system with advanced statistical methods. The evaluation studies so far performed were generally based on the assumption that errors follow a normal distribution. In fact, this hypothesis should be carefully verified in advance, in order to test how well the Gaussian classic statistics can adapt to data sets that are usually affected by asymmetrical gross errors. In those cases where even a proper process of outlier removal is not effective, robust and non-parametric methods for the derivation of accuracy measures should be preferred. This strategy can be further improved if the signed errors are computed, in order to analyse not only the dispersion values (which are related to the sensor noise), but also the bias effects on the object surfaces. Aim of the present paper is to test a rigorous statistical workflow for the evaluation of MMS data collected in a historic city centre. The site is selected in order to assess the system performance in a challenging urban scenario, that includes several building typologies with complex and diverse geometries/materials. As reference datasets we employ point clouds derived from a photogrammetric and a Terrestrial Laser Scanner (TLS) survey. Where possible, an error budget computation is discussed, in order to develop an adequate understanding of the most important components of uncertainty affecting the measurements.

2. PROJECT DESCRIPTION

2.1 MMS data with RIEGL VMX-450

The MMS RIEGL platform (Figure 1 and Table 1) integrates two synchronously operated VQ-450 laser scanners, a portable control unit (VMX-450-CU) and IMU/GNSS navigation hardware. The system is able to measure up to 1.1 million points and 400 profiles per second, providing extremely dense and feature-rich data even at high driving speed. Furthermore, it exploits the RIEGL echo signal digitization technology and online waveform processing, resulting in a high penetration

capability of obstructions (e.g. fences and vegetation). The platform is also equipped with the modular VMX-450-CS6 camera system, that complements the acquisition of geometric data with the recording of time-stamped images. Up to six industrial digital colour cameras with electronic shutters and 5 mm lens can be integrated. Camera calibration parameters are provided by the manufacturer, whereas exterior orientation parameters (i.e. time-stamped position and orientation of the cameras) are computed by RIEGL software in post-processing. The software includes several modules to manage each step of the processing pipeline, i.e. data acquisition (RiACQUIRE), data adjustment (RiPROCESS) and data geo-referencing (RiWORLD).



Figure 1. Configuration of the RIEGL VMX-450 system.

Sensor	VQ-450
Measuring principle	Time of Flight
Laser wavelength	Near infrared
Laser measurement rate	300 – 1100 kHz
Maximum range	140 – 800 m
Minimum range	1.5 m
Accuracy	8 mm, 1 σ
Precision	5 mm, 1 σ
Sensor	IMU/GNSS
Absolute position	0.020 – 0.050 m
Roll and pitch	0.005 °
True heading	0.015 °
Sensor	VMX-450-CS6
Resolution	5 Mpx
Sensor size	2452 x 2056 px
Pixel size	3.45 μ m
Nominal focal length	5 mm

Table 1. Technical characteristics of the RIEGL VMX-450 system according to the manufacturer.

2.2 Area of study and data collection

The data acquisition took place on September 22nd, 2014. An area of about 700 m west to east by 500 m north to south was covered in the city centre of Trento (Italy). It represents a typical urban scenario, that can be divided into two test-sites:

- the Duomo square (ca. 80 m x 70 m, Figure 2) surrounded by some of the most outstanding historical buildings of the city. Among them, five houses were selected for further analyses and segmented from the collected data. They include the Cathedral of San Vigilio in typical Romanesque style, the Medieval Palazzo Pretorio and some Renaissance buildings characterized by precious frescoed façades.
- narrow streets (average width of 3-4 m) around the Duomo square, surrounded by painted private and public buildings. Among them, three houses were identified and further investigated, including the backside façade of Palazzo Pretorio, Santa Annunziata church and the old municipal building of the city.

The selected areas (Figure 3) represent a challenging test-field for MMS mapping, whose metric performance depends mostly on the accurate determination of the time-variant position and orientation parameters of the mobile platform. Therefore, multipath and signal loss effects caused by tall buildings on narrow spaces can deteriorate the GNSS conditions and, consequently the trajectory computation. The tight urban canyons of the surveyed city centre are a typical example of these unfavourable scenarios.

Since the methodology adopted by the study is aimed at validating the final “commercial” product, the collected data was internally-processed by RIEGL experts and finally delivered as point clouds (LAS file format) in WGS84 global coordinate system. No attempts to evaluate each step of the calibration and adjustment processes are thus here undertaken. The point clouds are characterized by a mean spatial resolution of 2-3 cm and were exported separately for each of the two mounted laser scanners. Furthermore, three passes were collected in the Duomo square, resulting in six separate point clouds. This provided the possibility to assess the internal precision and consistency of the captured datasets.



Figure 2. Close-up view of the coloured 3D point cloud collected by MMS in the Duomo square of Trento (Italy).

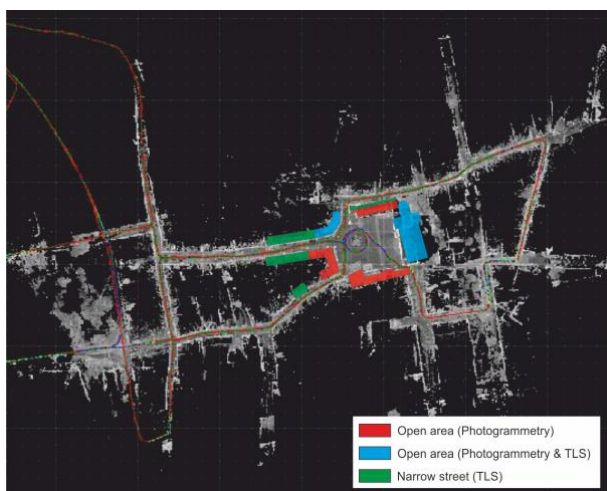


Figure 3. The trajectory of the area surveyed with the MMS and the buildings selected for statistical analyses in the square (red and blue symbols) and in the narrow streets (green symbols).

2.3 Reference data

2.3.1 Photogrammetric survey

The photogrammetric survey was carried out in the open square where the most notable historic buildings are located. This allowed an optimal design of the photogrammetric network,

whose layout was selected in order to optimize the accuracy and completeness of the final 3D point cloud. Only the buildings in the square (red and blue symbols in Figure 3) were thus reconstructed.

The image acquisition phase was performed using a Nikon D3X digital camera (35.9×24.0 mm sensor size, $5.95 \mu\text{m}$ pixel size) equipped with two fixed focal length lenses (35 and 50 mm). A total of 359 images were collected at a mean camera-object distance of 25 m, achieving an average GSD of 3.5 mm. Using design equations and an *a-priori* image measurement accuracy of 1/2 pixel (Luhmann et al., 2006), the range uncertainty and lateral resolution on the final model were anticipated to be in the [4.0;4.2] mm and [2.9;4.2] mm ranges, respectively.

Both image orientation and dense image matching were carried out using Agisoft Photoscan. 18 natural GCPs, measured with a total station, were used as observed unknowns in the bundle adjustment for the image orientation phase. The re-projection error was 0.52 pixels. In order to achieve a reasonable trade-off between processing effort and resolution, the successive dense matching was performed using the second-level image pyramid. Thus the derived dense point cloud (Figure 4) features a mean spatial resolution of 7 mm and consists of more than 107 million points.



Figure 4. Photogrammetric dense reconstruction of the Duomo square in Trento (Italy).

2.3.2 Terrestrial Laser Scanner survey

A Leica HDS7000 TOF CW laser scanner was used to acquire ultra-dense point clouds of two symmetrical buildings in the square (Figure 5 - blue blocks in Figure 3) and of those in the narrow streets (green blocks in Figure 3). More than 15 different scans were necessary to achieve a complete 3D reconstruction of the scene, with a mean instrument-object distance of 10 m.

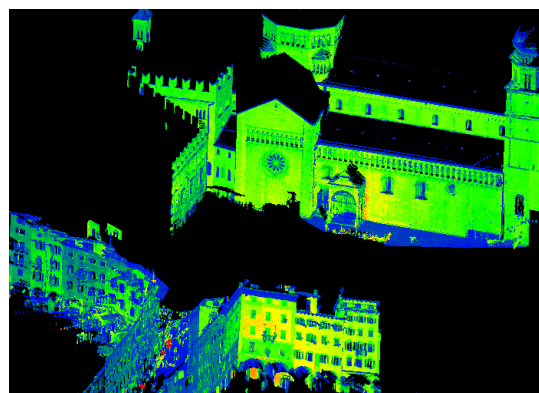


Figure 5. TLS ultra-dense reconstruction of two symmetrical buildings in the Duomo square of Trento (Italy).

Adopting the error propagation formulated in (Reshetyuk, 2009) and computing the beam divergence according to (Lichti and Gordon, 2004), the *a-priori* error on the surface is expected to be 3 mm. The alignment of the TLS scans in a local coordinate system was carried out using both planar contrast targets and natural points. No absolute geo-referencing was performed. The final registration error was below 4 mm. This last value can be considered a more realistic representation of the noise level present in the final 3D point clouds. The final point cloud is characterized by a mean spatial resolution of 3 mm, resulting in more than 240 million points.

3. METHODOLOGY

3.1 Cloud-to-Cloud signed distance computation

In order to perform the statistical analyses with a reasonable computational effort and to highlight any effects dependent on the building typology, both MMS and reference data are segmented in separate blocks. The Duomo square is divided into five blocks (hereinafter termed “Cathedral”, “Palace”, “Café”, “Houses” and “Hotel” in accordance with their main present-day function), whereas in narrow streets three distinct buildings are segmented (hereinafter called “Garibaldi”, “Belenzani” and “Cavour” from the name of the corresponding street). After the removal of non-significant points (i.e. those belonging to time-variant elements such as pedestrians and vegetation), the MMS data is registered with the reference ones (both photogrammetric and TLS point clouds) using the ICP algorithm implemented in Trimble Realworks. For each segmented building, cloud-to-cloud signed distances are then computed adopting the method Multiscale Model to Model Cloud Comparison (Lague et al., 2013). This algorithm, implemented in CloudCompare v2.6, performs a direct comparison of point clouds in 3D, avoiding the preliminary phase of meshing or gridding. Two steps are sequentially performed. At first, surface normals are estimated and oriented in 3D at a user-selectable scale, that should be consistent with the local surface roughness. Then, for each selected core point, the local distance between the two clouds is extracted. A confidence interval depending on point cloud roughness and registration error is computed as well. In the presented tests, each point of the MMS data is selected as “core” and a 50 cm normal scale is adopted in accordance with the local spatial resolution and suggestions given in Lague et al. (2013).

3.2 Statistical analyses

Our accuracy tests are performed using ‘reference’ data derived from a photogrammetric and a TLS survey. When comparing point clouds acquired by different instruments, some issues should be considered:

- a proper error budget computation should be defined, since each system has its own sources of uncertainties and sensitivity to outlier presence.
- the compared 3D points are not exactly corresponding to each other.
- the object surfaces are not equally digitized as the acquisition positions are different.

Due to these reasons, the number of gross errors affecting the datasets increases, thus requiring a rigorous statistical approach to be applied. The workflow adopted in this study is summarized in Figure 6 and follows Rodríguez-González et al. (2014).

In order to prevent possible bias effects affecting the data processing, the normality assumption, i.e. the hypothesis that

errors follow a Gaussian distribution, should be carefully checked. This can be carried out in various ways. Among them, the most common method is represented by the normality tests (D’Agostino, 1986). However, this approach does not work properly with very large samples (Rodríguez-González et al., 2014). In that case, a well-suited diagnostic should be preferred, such as the study of the graphical deviation in the Q-Q plot (Hasan et al., 2011; Hoehle, 2011). It reports the quantiles of the empirical distribution function with respect to the theoretical quantiles of the normal distribution function. If errors are distributed according to the Gaussian function, the Q-Q plot should be a diagonal straight line. Additional parameters, such as Skewness and Kurtosis can further support the normality assessment. The Skewness provides an indication of departure from symmetry in a distribution (asymmetry around the mean value) and is expressed as:

$$\text{Skewness} = \frac{\sum_{i=1}^n (x_i - \mu)^3}{(n-1)\sigma^3} \quad (1)$$

The Kurtosis is a measure of whether the data are peaked or flat relative to a normal distribution and is expressed as:

$$\text{Kurtosis} = \frac{\sum_{i=1}^n (x_i - \mu)^4}{(n-1)\sigma^4} \quad (2)$$

If the distribution is perfectly normal, Skewness and Kurtosis values of zero and three are obtained, respectively.

If the sample follows a Gaussian distribution, the error population can be characterized using the mean error (μ) and the standard deviation (σ). On the other hand, if the sample is not normally distributed, either due to the presence of outliers or because a different population’s hypothesis applies, a robust model based on non-parametric estimation should be employed. In this case, the median (m) and the median absolute deviation (MAD) are used as robust measures instead of the mean and standard deviation, respectively. The MAD is defined as the median (m) of the absolute deviations from the data’s median (m_x):

$$\text{MAD} = m(|x_i - m_x|) \quad (3)$$

In this work, both Gaussian statistics and the robust ones are computed for each comparison. Furthermore, analyses are performed not only on the raw data, but also on those derived from a filtering process. The latter is performed by applying the classical 3σ rule for outliers removal (Nkurunziza and Vermeire, 2014).

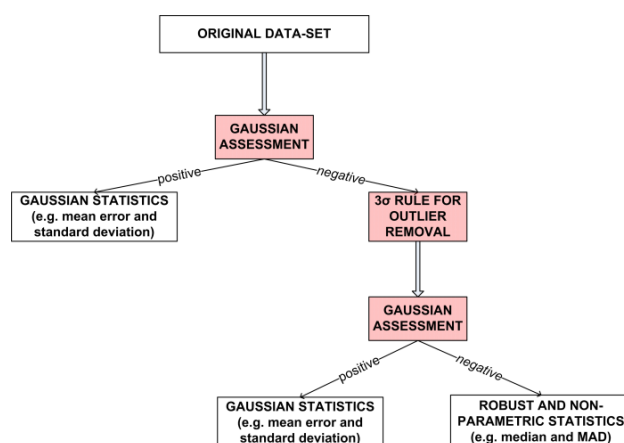


Figure 6. Statistical workflow.

4. RESULTS

Following the methodology described in Section 3, a two-folds study is then carried out assessing precision and accuracy of the Trento's MMS data. The precision of the MMS point clouds is a measure of the scatter of observed values whereas the accuracy is meant as the agreement between measurement results and the corresponding true value.

The MMS data precision is determined by comparing the different scans captured by the two lasers scanners at the same time (i.e. same timestamps) and in two subsequent acquisitions (i.e. different timestamps) on the same common surface. The Cathedral façade is selected as test-area for this analysis. Of course, for perfect geo-referencing and calibration procedures, no discrepancies between the measures from different scanners and/or different epochs should be present. In fact, this problem may arise, as also previously revealed (Barber et al., 2008; Haala et al., 2008). Furthermore, several passes are usually required in order to achieve a complete reconstruction of complex buildings, especially when MMS mapping is applied to historic city centres. The quality of the post-processing adjustment should be thus assessed.

On the other hand, the MMS data accuracy is evaluated. Reference measures are derived from the photogrammetric and TLS surveys. The former is restricted to the buildings in the open square, where the photogrammetric results could be optimized due to a favourable camera network. On the other hand, TLS is exploited for the acquisition of ultra-dense point clouds in the narrow streets, where the GNSS quality is reduced and MMS acquisition angles are far from the optimal case. TLS measurements are acquired in the square too, in order to achieve a denser "ground truth". However, this is only feasible for two selected areas, due to the considerable effort for data collection.

4.1 Precision evaluation

Among the three subsequent passes collected in the square, two acquisition stripes are selected for the precision evaluation, since they both allow a good coverage of the Cathedral façade. At each of the two epochs (denoted as t_0 and t_1), two point clouds are recorded by the two laser scanners on the MMS (named LS_1 and LS_2). Four separate point clouds are thus recorded and compared. Data acquired by LS_1 at the epoch t_0 (named LS_{1-t_0}) are selected as reference for all comparisons. Following the workflow outlined in Figure 6 and given the large size of the samples, the Gaussian assessment is performed by evaluating the graphical deviation in the Q-Q plot. For each of the three comparisons, the statistical graphic is derived for the error population computed with the Multiscale Model to Model Cloud Comparison. The original distance dataset is then filtered with the 3σ criterion and the corresponding Q-Q plot is extracted again. Figure 7 shows the plots obtained from the comparison of LS_{2-t_0} point cloud against the reference one (LS_{1-t_0}). The Q-Q plot shape extracted for the original (raw) dataset (Figure 7a) is significantly far away from the ideal normal hypothesis. On the other hand, the error sample derived from the outlier filtering is closer to the Gaussian distribution (Figure 7b), but not enough to allow a parametric error estimation. Both the parametric and non-parametric statistics are finally computed and summarized in Table 2.

As expected, Kurtosis and Skewness parameters agree with the outcomes of the graphical analyses (Q-Q plots): a significant percentage of outliers affects the original population and should be removed. Furthermore, by considering the robust error measures computed after the outlier removal, a mean bias between point clouds of 2.6 mm is observed. On the other hand, the robust measure of dispersion (MAD) reaches up to

± 5.8 mm, and increases up to ± 8.6 mm under a Gaussian hypothesis¹. If comparisons performed with data acquired at the same epoch provide a means of assessment for errors related to LS noise and calibration uncertainty, the analyses on multi-epochs data account also for the post-processing error propagation. Values of median and MAD increase when point clouds with different timestamps are compared, highlighting some expected limitations of the geo-referencing procedure. The latter, in fact, was performed without including any ground control information within the trajectory adjustment procedure. Finally, along with the numerical statistics, a spatial distribution of the discrepancies can be studied. The colour-coded maps of signed distances reveal the presence of typical error stripes, that are more or less coincident with scanning orientation. For instance, Figure 8 shows the errors computed from the comparison of LS_{2-t_0} point cloud against the reference one. The scale ranges from -0.01 m (blue) to 0.01 m (red).

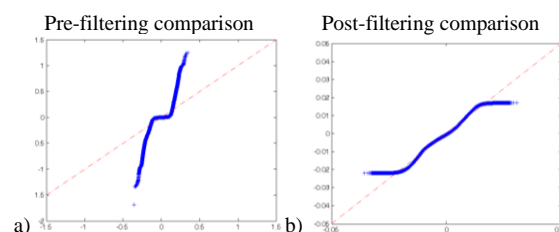


Figure 7. Q-Q plots of the comparison LS_{2-t_0} vs. LS_{1-t_0} . The quantiles of input sample (vertical axis) are plotted against the standard normal quantiles (horizontal axis). Pre-filtering (a) and post-filtering (b) data are shown.

Raw comparison	LS_{2-t_0}	LS_{1-t_1}	LS_{2-t_1}
Number of points	3,651,128	829,741	661,362
Sample mean (cm)	-0.78	-0.06	-0.13
Standard deviation (cm)	± 6.83	± 4.96	± 4.92
Median (cm)	-0.27	-0.25	-0.26
MAD (cm)	± 0.43	± 0.62	± 0.65
Kurtosis (-)	47.24	69.59	75.46
Skewness (-)	-0.25	-0.01	0.57
Filtered comparison	LS_{2-t_0}	LS_{1-t_1}	LS_{2-t_1}
Number of points	3,126,270	738,842	581,103
Sample mean (cm)	-0.25	-0.28	-0.36
Standard deviation (cm)	± 0.65	± 0.93	± 0.89
Median (cm)	-0.24	-0.27	-0.28
MAD (cm)	± 0.35	± 0.54	± 0.58
Kurtosis (-)	3.74	3.60	3.34
Skewness (-)	-0.08	0.00	-0.09

Table 2. Statistical results of the raw and post-processed MMS data analyses.

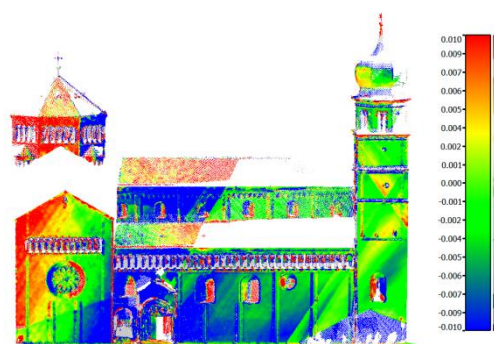


Figure 8. Colour-coded map of the comparison between LS_{2-t_0} vs. LS_{1-t_0} . Differences range between -0.01 m (blue) and 0.01 m (red). White areas represent no-data surfaces.

¹ The Gaussian estimation of the MAD is computed as $(1.4826 \cdot \text{MAD})$

4.2 Accuracy evaluation

To evaluate the accuracy of a single MMS point cloud against the reference data (photogrammetry and TLS), a pre-registration phase with ICP is carried out, in order to dismiss the uncertainties associated to the global coordinate geo-referencing. Only the errors affecting the 3D modelling and reconstruction of the architectural element are thus analysed.

4.2.1 Comparison with the photogrammetric reference data

All Q-Q plots prove that distances computed with the raw datasets are not suitable for a Gaussian estimation. As shown from the example given in Figure 9a, the input samples are far away from the ideal distribution. Gross errors should thus be filtered out, since they can corrupt the true statistical distribution of discrepancies. This evidence is verified again by the extremely high values of Kurtosis and Skewness parameters (see Table 3 - upper part).

After the filtering, the resulting datasets are more close to a normal distribution: this is confirmed by both graphical analyses (Q-Q plot in Figure 9b) and numerical ones (Kurtosis and Skewness parameters in Table 3 - bottom part).

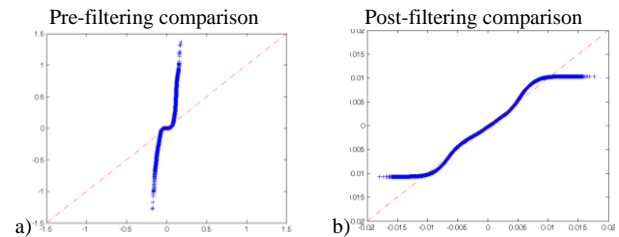


Figure 9. Q-Q plots of the comparison MMS vs. Photogrammetry for the Café dataset. The quantiles of the input sample (vertical axis) are plotted against the standard normal quantiles (horizontal axis). Pre-filtering (a) and post-filtering (b) data are shown.

However, a robust analysis of the MMS accuracy is here advised. Table 3 (bottom part) lists the derived statistical values, that allow a robust estimation of the error. This can be defined as median \pm MAD and it is here estimated as $-0.7 \text{ mm} \pm 4.0 \text{ mm}$. If a Gaussian measure of the MAD is computed, the error dispersion grows up to $\pm 5.9 \text{ mm}$. This second value can be compared against the *a-priori* value of uncertainty quoted for the reference dataset: results are consistent with both the photogrammetric lateral resolution and range uncertainty.

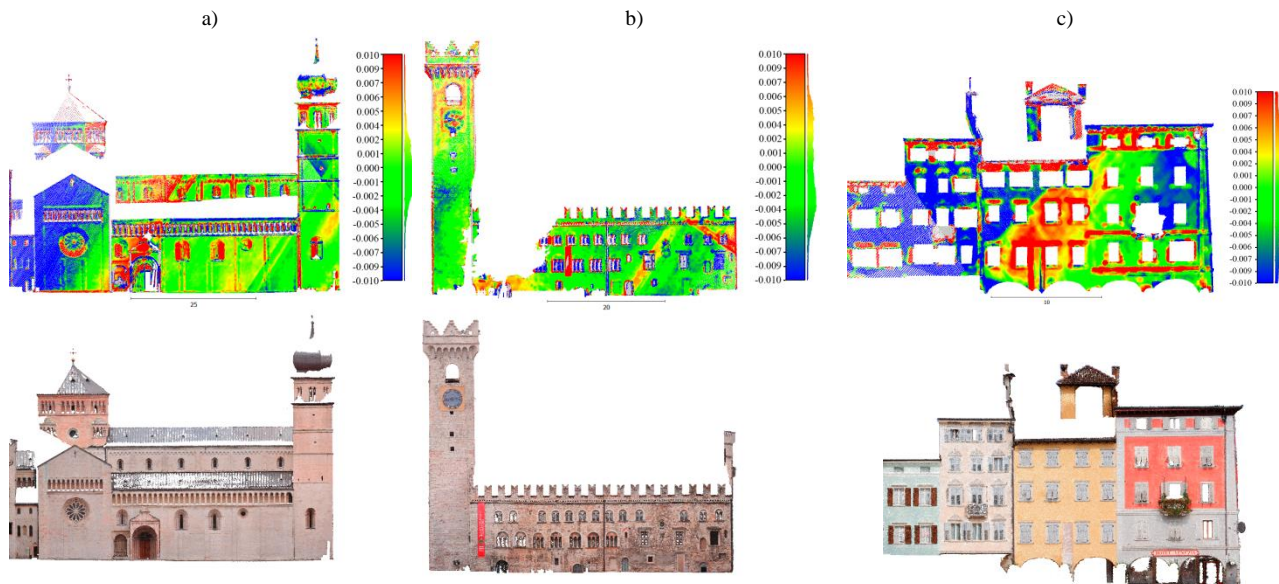


Figure 10. Colour-coded maps of the comparison analysis between the MMS and Photogrammetry data for the Cathedral (a), Palace (b) and Hotel (c) datasets. The scale ranges from -0.01 m (blue) to 0.01 m (red). Corresponding photogrammetric point clouds are shown below.

Raw comparison	Café	Cathedral	Palace	Houses	Hotel
Number of points	249,368	3,257,278	687,885	1,372,294	695,607
Sample mean (cm)	-0.69	-0.12	0.25	-0.40	0.28
Standard deviation (cm)	± 6.42	± 3.51	± 3.95	± 8.68	± 3.86
Median (cm)	-0.03	-0.03	0.07	-0.36	0.09
MAD (cm)	± 0.26	± 0.25	± 0.34	± 0.92	± 0.60
Kurtosis (-)	48.78	118.65	65.49	30.01	84.76
Skewness (-)	-1.77	-0.05	1.62	1.06	3.21
Filtered comparison	Café	Cathedral	Palace	Houses	Hotel
Number of points	192,385	2,716,260	538,369	1,086,039	644,286
Sample mean (cm)	-0.04	-0.02	0.05	-0.32	0.06
Standard deviation (cm)	± 0.31	± 0.35	± 0.44	± 1.17	± 0.91
Median (cm)	0.00	-0.04	0.08	-0.31	0.07
MAD (cm)	± 0.19	± 0.20	± 0.26	± 0.71	± 0.55
Kurtosis (-)	3.39	3.79	3.42	3.67	3.31
Skewness (-)	-0.27	0.26	-0.23	-0.08	0.08

Table 3. Statistical results of the accuracy evaluation: MMS vs. Photogrammetry comparisons on the raw and filtered point clouds.

Finally, along with the numerical error assessment, the spatial error distribution is highly relevant for assessing the MMS performance in modelling architectural façades. For this reason, three colour-coded maps of signed distances are showed in Figure 10, where the scale ranges from - 0.01 m (blue) to 0.01 m (red).

The Cathedral dataset covers a large area: 65 m in width and up to 50 m in height. The corresponding error distribution (Figure 10a) shows that the regions delivering the largest deviations are:

- areas acquired from high distances and with high acquisition angles, such as the left portion of the façade and the upper part of the tower (see also the trajectory in Figure 3);
- areas characterized by challenging materials (e.g. windows) and complex geometry (e.g. the main entrance).

Furthermore the same error patterns, previously highlighted by the MMS-MMS comparisons (see Figure 8), are here present too. The diagonal stripes may be linked to the direction of the scanning acquisition or probably due to internal calibration errors. These bands are characterized by a rapid change of the errors, that may cause a degradation of the final 3D reconstruction.

In the Palace dataset, beside a lack of data in the upper left corner of the building due to the presence of a large tree, systematic error stripes related to the LS orientation are found, together with an error increment in the tower (Figure 10b). The latter effect is again derived from the increased acquisition distance and scanning angle.

Finally, an area of buildings characterized by different materials and a less detailed surface is analysed (Café, Houses and Hotel datasets). The bottom part is removed from the investigation due to the presence of obstructions, such as cars, vegetation, people, etc. In addition, all time-variant elements are removed as well. For instance, a small area of 33 m x 15 m is showed in Figure 10c, corresponding to the Hotel dataset. Higher errors are found, as it is confirmed by the robust error dispersion (MAD), which reaches up to ± 7.1 mm for the Houses dataset (Table 3). The causes of this behaviour are under study and may be related to the material properties.

4.2.2 Comparison with the TLS reference data

Graphical analyses carried out with the point clouds acquired in the Duomo square (Cathedral - Figure 11 - and Café datasets) prove that the Gaussian model doesn't fit well with the acquired MMS raw data. Also in this case a filtering is necessary and a robust analysis of the MMS accuracy is then performed. The statistics reported in Table 4 agree with results achieved by adopting the photogrammetric model as reference.

Raw comparison	Cathedral	Café
Number of points	2,437,121	155,470
Sample mean (cm)	± 1.68	± 0.40
Standard deviation (cm)	8.76	8.26
Median (cm)	± 0.21	± 0.22
MAD (cm)	0.41	0.34
Kurtosis (-)	46.12	25.44
Skewness (-)	5.33	-0.84
Filtered comparison	Cathedral	Café
Number of points	1,876,985	122,113
Sample mean (cm)	0.12	0.18
Standard deviation (cm)	± 0.48	± 0.35
Median (cm)	0.17	0.18
MAD (cm)	± 0.26	± 0.23
Kurtosis (-)	3.73	2.80
Skewness (-)	-0.01	-0.20

Table 4. Statistical results of the accuracy evaluation: MMS vs. TLS for the square dataset on the raw and filtered point clouds.

Raw comparison	Garibaldi	Belenzani
Number of points	4,853,987	1,156,009
Sample mean (cm)	-0.24	-0.15
Standard deviation (cm)	± 5.03	± 5.56
Median (cm)	-0.12	-0.52
MAD (cm)	± 0.28	± 0.95
Kurtosis (-)	77.99	50.58
Skewness (-)	-0.66	2.82
Filtered comparison	Garibaldi	Belenzani
Number of points	3,995,458	989,354
Sample mean (cm)	-0.13	-0.42
Standard deviation (cm)	± 0.43	± 1.26
Median (cm)	-0.10	-0.53
MAD (cm)	± 0.21	± 0.79
Kurtosis (-)	4.08	3.38
Skewness (-)	-0.03	0.07

Table 5. Statistical results of the accuracy evaluation: MMS vs. TLS for the street dataset on the raw and filtered point clouds.

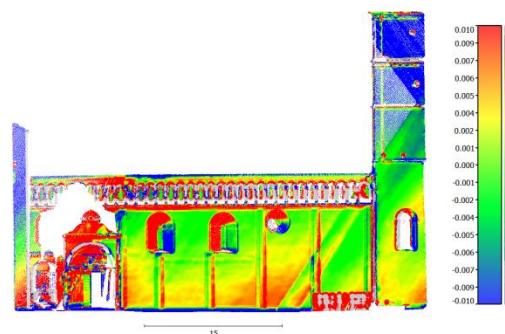


Figure 11. Colour-coded map of the comparison analysis between MMS and TLS data in the Cathedral dataset. The scale ranges from - 0.01 m (blue) to 0.01 m (red).

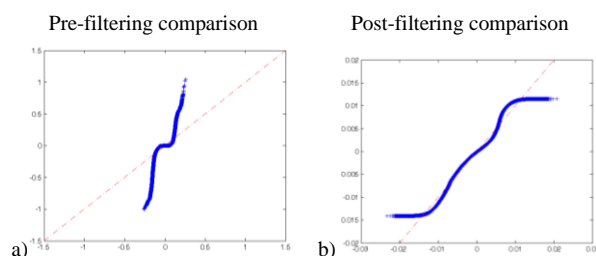


Figure 12. Q-Q plots of the comparison MMS vs. TLS for the Garibaldi dataset. The quantiles of input sample (vertical axis) are plotted against the standard normal quantiles (horizontal axis). Pre-filtering (a) and post-filtering (b) data are shown.

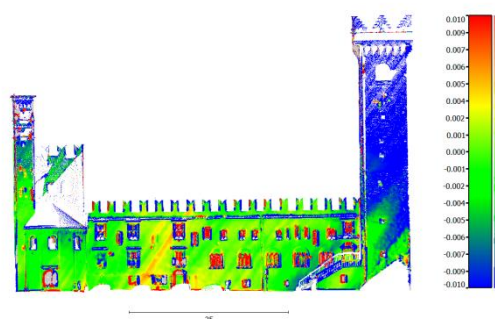


Figure 13. Colour-coded map of the comparison analysis between MMS and TLS data in the Garibaldi dataset. The scale ranges from - 0.01 m (blue) to 0.01 m (red).

The study is then completed by analysing the signed distances computed for the three façades on the narrow streets (Garibaldi, Cavour and Belenzani datasets). Due to the limited street width and significant height of the buildings, the acquisition angles are expected to be more severe here, deteriorating the metric accuracy of the resulting point clouds. Likewise all previous comparisons, Q-Q plots show a significant departure from normality, that is reduced after the 3σ filtering (Figure 12).

Statistics are listed in Table 5; results achieved by the Cavour dataset are here omitted, since the significant amount of occlusions doesn't allow a proper analysis. An average robust dispersion (MAD) of ± 5.0 mm is outlined, that reaches ± 7.4 mm if a Gaussian estimation is derived. As expected, this value is higher than the one delivered by comparisons of the square datasets, but the gap is not statistically significant.

Worth to note is finally the spatial error distribution, shown as colour-coded map of distances. For the Garibaldi dataset (Figure 13), it is still affected by the systematic diagonal stripes. Furthermore, the error increases quickly on the tower, whose height requires an unfavourable acquisition direction.

5. CONCLUSIONS

This paper strives to provide advanced statistical methods specifically designed to assess precision and accuracy of a MMS solution. The inter-comparison methodology here adopted is based on reference 3D imaging techniques (both active and passive sensors), whose uncertainty is predicted in advance, using manufacturer specifications and design equations. The MMS point clouds (Figure 14 – Duomo square dataset) are then compared against these reference ones and a statistical analysis of the resulting error distribution is performed.

The suggested approach relies on a Gaussian assessment, followed by an outlier filtering process. Finally, non-parametric statistical models are applied, in order to achieve a robust estimation of the error dispersion. All datasets show a large lack of symmetry that leads to the conclusion that the standard normal parameters are not adequate to assess this type of data. The use of non-normal statistics gives thus a more appropriate description of the error population and yields results that meet what may be expected (*a-priori* errors) concerning the assessment of both precision and accuracy.

The comparisons between point clouds acquired by the two laser scanners at the same epoch deliver a Gaussian estimation of the MAD in the order of ± 5.2 mm, which is consistent with the precision (1σ value) quoted by the MMS datasheet. The error increases up to ± 8.6 mm if the acquisitions in two different epochs are compared. These differences can be traced back to geo-referencing errors and surface-related problems. The former may be mitigated by including control point information within the post-processing adjustment. Depending on the final metric requirements, this issue should be adequately dealt with, especially if the MMS is used for accurately mapping historic city centres. Indeed this requires complex trajectories and multiple passes, thus demanding an optimization of the geo-referencing process.

The accuracy assessment of one single point cloud highlights the good metric potentialities of the RIEGL mobile system. Errors estimated for the buildings in the Duomo square are characterized by an average dispersion of ± 5.9 mm (MMS vs. Photogrammetry) and ± 3.6 mm (MMS vs. TLS). These values are consistent with the uncertainty quoted by the MMS manufacturer (8 mm, 1σ). As expected, errors increase if MMS mapping is performed in narrow streets with high buildings and more challenging obstacles. However, the computed value of ± 7.4 mm in such areas still represents an acceptable trade-off.

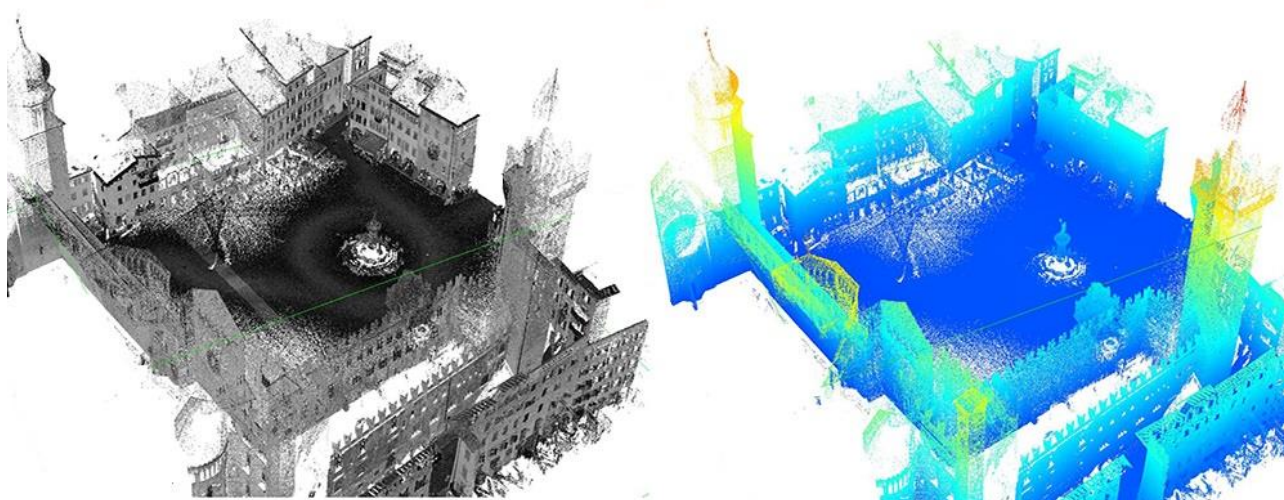


Figure 14. Intensity (left) and elevation (right) map of the MMS data collected in the Duomo square of Trento (Italy).

REFERENCES

Barber, D., Mills, J., Smith-Voysey, S., 2008. Geometric validation of a ground-based mobile laser scanning system. *ISPRS Journal of Photogrammetry and Remote Sensing*, 63(1), pp. 128–141.

Bitenc, M., Lindenbergh, R., Khoshelham, K., Van Waarden, A.P., 2011. Evaluation of a LiDAR land-based mobile mapping system for monitoring sandy coasts. *Remote Sensing*, 3(7), pp. 1472–1491.

Bossler, J.D., Novak, K., 1993. Mobile Mapping System: new tools for the data collection of GIS information. *Canadian conference on GIS'93*, Ottawa, Canada, pp. 306–315.

- D'Agostino, R.B., 1986. *Tests for the normal distribution*. Marcel Dekker.
- El-Sheimy, N., 1996. A mobile multi-sensor system for GIS applications in urban centers. In: *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, Vienna, Austria, Vol. XXXI, Part B2, pp. 95-100.
- El-Sheimy, N., Schwarz K.P., Gravel, M., 1995. Mobile 3-D positioning using GPS/INS/Video cameras. *The Mobile Mapping Symposium*, Ohio State, USA, pp. 236-249.
- Haala, N., Peter, M., Kremer, J., Hunter, G., 2008. Mobile LiDAR mapping for 3D point cloud collection in urban areas – a performance test. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, Rostock, Germany, Vol. XXXVII, Part B5, pp. 1119-1124.
- Hasan, A., Pilesjö, P., Persson, A., 2011. The use of LiDAR as a data source for digital elevation models – a study of the relationship between the accuracy of digital elevation models and topographical attributes in northern peatlands. *Hydrology and earth system sciences discussions*, 8, pp. 5497-5522.
- Hoehle, J., 2011. The assessment of the absolute planimetric accuracy of airborne laserscanning. In: *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, Calgary, Canada, Vol. XXXVIII-5W12, pp. 145-150.
- Kaartinen, H., Hyypä, J., Kukko, A., Jaakkola, A., Hyypä, H., 2012. Benchmarking the performance of mobile laser scanning systems using a permanent test field. *Sensors*, 12(9), pp. 12814-12835.
- Kawashima, K., Kanai, S., Date, H., 2012. Automatic recognition of piping system from laser-scale terrestrial laser scanner point cloud. *Journal of the Japan Society for Precision Engineering*, 78(8), pp. 722-729.
- Kremer J., Hunter G., 2007. Performance of the StreetMapper mobile LiDAR mapping system in real world projects. *Photogrammetric week 2007*, Stuttgart, Germany, pp. 215-225.
- Lague, D., Brodu, N., Leroux, J., 2013. Accurate 3D comparison of complex topography with terrestrial laser scanner: Application to the Rangitikei canyon (N-Z). *ISPRS Journal of Photogrammetry and Remote Sensing*, 82, pp. 10-26.
- Lichti, D., Gordon, S., 2004. Error propagation in directly georeferenced terrestrial laser scanner point clouds for cultural heritage recording. *FIG working week*, Athens, Greece.
- Lim, S., Thatcher, C. A., Brock, J. C., Kimbrow, D. R., Danielson, J. J., Reynolds, B. J., 2013. Accuracy assessment of a mobile terrestrial LiDAR survey at Padre Island National seashore. *ISPRS Journal of Photogrammetry and Remote Sensing*, 34(18), pp. 6355-6366.
- Luhmann, T., Robson, S., Kyle, S., Harley, I., 2006. *Close range photogrammetry: Principles, methods and applications*. Whittles, Dunbeath, UK, 500 pages.
- Nkurunziza, M., Vermeire, L., 2014. A comparison of outlier labeling criteria in univariate. *9th international conference on teaching statistics*, Flagstaff, Arizona (USA).
- Pu, S., Rutzinger, M., Vosselman, G., Elberink, S., 2011. Recognizing basic structures from mobile laser scanning data for road inventory studies. *ISPRS Journal of Photogrammetry and Remote Sensing*, 66(6), pp. S28-S39.
- Puente, I., González-Jorge, H., Martínez-Sánchez, J., Arias, P., 2013. Review of mobile mapping and surveying technologies. *Measurement*, 46(7), pp. 2127-2145.
- Reshetyuk, Y., 2009. *Self-calibration and direct georeferencing in terrestrial laser scanning*. PhD thesis, KTH - Royal institute of technology, Stockholm, Sweden.
- Rodríguez-Gonzálvez, P., Garcia-Gago, J., Gomez-Lahoz, J., González-Aguilera, D., 2014. Confronting passive and active sensors with non-gaussian statistics. *Sensors*, 14(8), pp. 13759-13777.
- Slattery, K., Slattery, D., Peterson, J., 2012. Road construction earthwork volume calculation using three dimensional laser scanning. *Journal of Surveying Engineering*, 138(2), pp. 96-99.
- Tao, V., Li, J., 2007. *Advances in mobile mapping technology*. ISPRS Book Series No. 4; Taylor & Francis: London, UK, 176 pages.
- Ziparo, V.A., Zaratti, M., Grisetti, G., Bonanni, T.M., Serafin, J., Di Ciccio, M., Proesmans, M., Van Gool, L., Vysotska, O., Bogoslavskyi, I., Stachniss, C., 2013. Exploration and mapping of catacombs with mobile robots. *IEEE international symposium on safety, security, and rescue robotics (SSRR)*, pp. 21-26.