Comparing performance evaluation methods for assessing an industrial robot tracking system

O. Martin 1, N. Southon, 1, P. Meekums 1, C. Davey 1, N. Yu 2
[1]*Insphere Ltd, BS16 7FR, UK*
[2]*School of Engineering, University of Edinburgh, EH9 3FB, UK*

Abstract

Performance evaluation of an instrument is typically conducted to understand the system's capability in question and inform an end-user. However, the end-users intended application may differ in measurement requirements and environmental conditions from that of the performance evaluation.

Several standards begin to address the performance evaluation of 6DOF metrology systems, including VDI/VDE 2634-1/2/3, ISO 10360-13, ASTM E3064-16 and ASTM E2919-14. These standards have been constructed with a specific instrument paradigm in mind; this includes the instrument configuration, targeting, use case and industry.

IONA is an example of a robot tracking system for which end-users desire performance evaluation. It consists of a network of sensors with no fixed operational volume; additionally, IONA measures constellations of targets, not a single point in the traditional sense. As such, we can draw on the standards above for guidance but must modify them to ensure the test can accommodate the instrument's operational parameters and use case. However, even in their altered states, there is variation in the proposed testing procedures and analysis, giving rise to different expressions of performance. Notably, relative (inter-point) distance comparisons may not be sensitive to systemic bias evident when an analysis in absolute coordinates is carried out.

This paper demonstrates how different evaluation techniques can be applied to the same test data illustrating how bias can go undetected and performance metrics can vary. The test was carried out in an industrial setting to ensure it represents the end-users intended environment.

1 Introduction

This paper examines how different expressions of the same dataset can give rise to differing IONA system performance evaluation metrics. The evaluation techniques draw on but are not limited to, the relevant standards. The purpose of the test, analysis, and metrics – from an end-user perspective - is to understand how the metrology system performs in the intended operational environment and working volume. This real-world performance requirement is typically different from the lab-based performance evaluation setups giving rise to uncontrolled environmental conditions, sub-optimal instrument and target mounting locations and line-of-sight obstructions.

The measurement system detailed in the paper is the IONA system from INSPHERE. A robot tracking system consisting of three or more stereo-pair sensors distributed around the operational volume of the automation. The system measures clusters of retro-reflective spherical targets (known as constellations) to determine points of interest (in 6DOF) within the manufacturing environment.

Although striving for generality, performance verification standards have been designed around specific instruments and industries that do not cater to new and emerging metrology systems such as IONA. For example, IONA operates as a network of sensors with no fixed operational volume; additionally, IONA measures constellations of targets, not single points in the traditional sense. As such, we can draw on the standards for guidance but must modify them to ensure the test is fit for purpose and practical to carry out.

The cell set-up for the performance test was not optimised and is in a simulated industrial environment to ensure the test results give representative values for the expected performance in the real world.

2 Applicable Standards

Several applicable measurement standards begin to address the performance testing of 6DOF metrology systems. These include VDI/VDE 2634-1/2/3, ISO 10360-13, ASTM E3064-16 and ASTM E2919-14. Whilst they all aim to quantify the metrological performance of a system objectively, the choices in how to measure and evaluate system performance can influence the comparability of performance across standards.

2.1 VDI/VDE 2634-1/2/3

The VDI/VDE 2634 series has three parts. Part 1 [1] is concerned with point-to-point measurements, that is, the evaluation of an artefact with at least 5 test lengths arranged in specified orientations in the measurement volume. Part 2 [2] concerns area scanning systems with a single instrument viewpoint. Part 3 [3] is similar to Part 2 but considers the measurement system having multiple viewpoints, either from a single system taking multiple measurements around a part or multiple measurements being taken simultaneously by a network of sensors. VDI/VDE 2634-2 & -3 have been commonly used by surface-scanning photogrammetry and fringe projection-based measurement systems. The first sections concern the measurements of spheres (probing size and probing form errors). The following sections consist of ball bar measurements arranged in the same orientations as VDI/VDE 2634-1 (sphere spacing errors) and length measurement errors. IONA's operation and data collection do not neatly fit into any single part of the VDI/VDE 2634 standard – the system has a variable operational volume such as those described in part 3. However, it reports point-to-point measurements found in part 1. It is essential to remember that a single 6DOF measurement in IONA is derived from 3D measurements of a constellation of points; these 3D measurements are not exposed to the user.

2.2 ISO 10360-13

ISO 10360-13 [4] is a logical extension of the ISO 10360 series of standards to volumetric 3D measurement and echoes much of the VDI/VDE 2634 standard. It is written for a system that produces a 3D point cloud of viewed surfaces; the standard is focused on artefact-based references. The first sections concern sphere measurements (probing size and form errors). The second section contains ball bar and ball beam measurements (distortion and length measurement errors, respectively). Finally, it also contains a flat form measurement error section where a test flat has its surface measured in many poses throughout the measurement volume. IONA is not accommodated as it has a non-fixed operational volume and does not measure single points in the sense that is expected in the standard. The artefact would be very difficult to produce, commission and use in reality – as a minimum requirement for IONA would be a 2m+ long beam with large clusters of targets at each end.

2.3 ASTM E3064-16

ASTM E3064-16 [5] is a standard aimed at motion capture systems. It contains a ball beam measurement in a sweep of the volume instead of a set of static locations. The standard is very clearly optimised for motion capture systems. So some requirements, such as walking the calibrated ball bar through the volume at a walking pace, are not optimal for a manufacturing measurement system. However, ASTM E3064-16 does analyse all six degrees of freedom and allows for repeatability measurement evaluation without accuracy evaluation if reference measurements are not possible. It is also the only standard to calculate root mean square error, maximum error, and percentile error. The reference artefact is specified to be 300 mm – unlike the above standards that have a measurement proportional to the operational volume. In many respects, this standard has the closest alignment to IONA regarding expected system operation. However, the evaluation techniques do not align well with manufacturing engineering requirements. Motion capture systems are increasingly used in manufacturing challenges where large-volume measurement systems are typically deployed. These [systems] are expected to have a known specification throughout the operational volume and – in many cases – in absolute terms.

2.4 ASTM E2919-14

ASTM E2919-14 [6] is a standard for static 6DOF measurement systems. Unlike the other standards mentioned, it is not prescriptive about the spatial arrangement of the measurement procedure, opting instead for the 'random pose' choice of the user evaluating the system. System performance evaluation is achieved with statistical hypothesis testing across the data's mean, quartile, maximum and variation. It requests a reference measurement system with ten times smaller measurement uncertainty than the measurement uncertainty of the system under test, which is difficult for large-volume measurement systems like IONA. Critically, it checks both relative and absolute measurement performance. The reference measurement system might limit the potential test poses through the line of sight/operational constraints. This [constraint] may, in turn, interfere with the statistical assumptions of the performance evaluation.

2.5 Summary

Ball bar length measurements are a common feature in most of the standards. They are a common artefact for CMM measurements because individual measurement errors can be minimised through the averaging inherent in calculating the centre of the spheres and, therefore, their offset from each other. Because the ball bar measurements may rely on sphere measurements derived from the 3D surface measurement data, the VDI/VDE and ISO standards contain sections to isolate the sphere measuring the performance of each system. It is an interesting nuance that many metrology systems can comply with the standards by measuring features, whether a sphere in isolation or several spheres. This can improve the stated performance as a form of averaging occurs when constructing the geometry. This methodology also relies on the metrology system having a 'probing' operation which – as in the case of IONA – is not always available.

In the case of VDI/VDE and ISO standards, the length of the ball bar (or point-to-point) measurements are functions of the measurement volume. Artefact-based ball beams/bars become impractical when the length exceeds 2m; they are impractical to handle and characterise with a traceable system.

All standards quantify performance, at least to some extent, by Maximum Permissible Error (MPE). Since there is no standard for determining the MPE of a measurement system, we have quoted the maximum errors obtained during the analysis. However, if performance metrics such as MPE are desired in a measurement system, and none is stated, it can be evaluated with careful statistical analysis [7]. ASTM E2919-14 uses a statistical truncation check. All three standards have accommodations for filtering outlying data from the measurement population before comparing it against reference measurements. However, all the standards also require unfiltered results to be stated.

It is a common approach for the standards to attempt to characterise the system performance through a series of relative or local measurements; for example, sphere diameters and point-to-point distances. Although understandable for practical reasons – it is easier to parameterise a relative test for generality – many industries are interested in using measurement systems for absolute measurements. Absolute measurements have a global reference or datum from which all measurements are reported in a Cartesian format. A method of coordinate comparison can be carried out to determine the absolute performance of a system [8,9,10]. However, the only standard that employs this approach, ASTM E2919-14, does not specify the spatial constraints on which measurement poses should be used for a measurement system evaluation.

It is of note that the ASTM standards address the angular performance of the measurement system, accommodating 6DOF tracking systems such as IONA. However, creating a suitable reference standard with a low uncertainty is non-trivial – but was attempted during the data collection.

3 Dataset Collection

A single measurement from the IONA system reports the 6DOF transform between two measured frames, derived from the 3D measurement of numerous retro-reflective spheres. This is not a single-point to single-point measurement, nor a point cloud measurement in the traditional sense. IONA can measure many transforms in a single capture. The reporting of each measurement can be relative to a single global datum or a unique datum required for each reported transform.

The closest analogue to a ball bar in the IONA measurement paradigm would be a beam with a cluster of spheres at each end. This cluster would have to be of a representative scale to the constellations used in most IONA applications to ensure the evaluation is comparable to its operational use. These constellations would be cubes of at least 250mm in length sides – which is typical for a mid-sized robot end effector. This artefact would be challenging to construct in terms of stiffness and characterising the reference length. The reference length would need to span around 2/3 of the working volume (depending on the standard), which is not fixed and could be tens of metres. Therefore, a more practical solution is to create a virtual scale bar using a more accurate and established metrology system, for example, a laser tracker. The laser tracker will capture data in the same location and timeframe as the IONA system. The laser tracker has a stated uncertainty of:

±15µm + 6µm/m (2 sigma) [11].

The above equation gives us an expected uncertainty of 45µm at 5m. Testing IONA violates the rule of thumb that reference uncertainty should be less than 10% of the system under evaluation, as specified by ASTM E2919-14. In addition, the T-Mac will provide us with Rx, Ry, & Rz data. The ASTM standard is the only procedure to include angular pose measurement. Also, it only considers two small constellations that have a small displacement between each other. The IONA system works with much larger constellations and larger displacements between them, so the methodology needs adapting.

To accommodate the competing specifications within the standards - for the length and position of the ball - it was decided to capture a closely spaced (≈300 mm) grid of points within the operational volume of the automation platform, not the total measurement volume; this is a similar philosophy to the ISO 9283 [12] standard for performance testing of an industrial robot. This grid of data enables multiple analyses to be performed.

The IONA-measured grid and the laser tracker reference measured grid have a common global coordinate system to evaluate the absolute measurement performance. Each IONA-measured point within the grid can be compared in absolute terms or relative to another point within the grid. This allows for the evaluation of the maximum, average, and variation in measuring performance. Also, this grid tests the operational volume more thoroughly than the subset required by the standards.

3.1 Experimental Setup

A 6DOF laser tracker (Hexagon/Leica AT960 & T-Mac) was used as a reference standard in place of a scale bar (due to the magnitude of the test volume) to match the 6DOF nature of the IONA measurement of constellations.

The laser tracker and IONA reported measurements from a common frame of reference (the fixture datum). The T-Mac was attached to the robot end-effector along with a set of IONA targets (collectively a constellation) such that both measurement systems could acquire data simultaneously. ORA – the software for setting up and collecting IONA data - was configured so that the reported transform coincided with the native T-Mac reported frame. The setup can be seen in Figures 1 & 2.

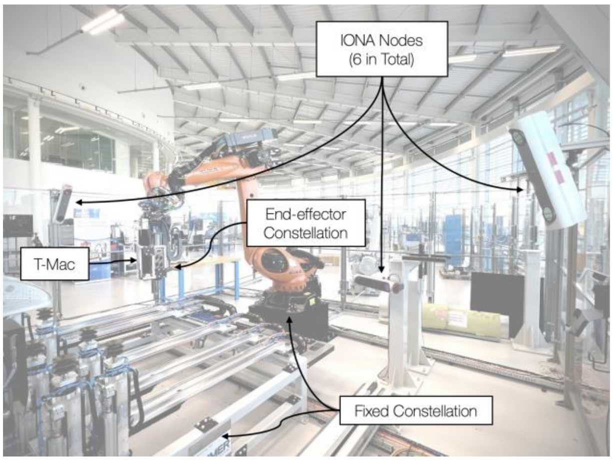

Figure 1: the experimental setup.

A grid of 287 points equispaced in a volume of 0.7 m x 1.4 m x 0.7 m was measured by both the laser tracker and IONA at the same time. The IONA volume is not fixed and is scalable. Subsequently, the volume was chosen to represent 2/3 of the typical operational volume of the robot in line with the VDI/VDE 2634-3 standard (Figure 2). The operational volume is 1m x 1m x 2m. The cell consisted of 6 IONA nodes. The orientation of the end-effector remained constant to ensure the laser tracker could maintain the line of sight. The robot was employed as a positioning device; however, it is important to note that robot errors do not contribute to the test results.

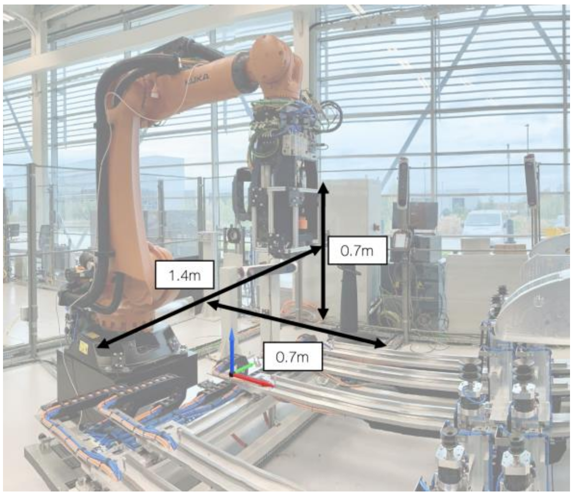

Figure 2: the testing volume

4 Results

The results can be analysed and expressed in several ways:

1. Inter-point distance of all measurements: analogous to ball beam measurements & aligned to the VDI/VDE 2634-3 & ISO 10360-13 standards.
2. measurements: analogous to ball bar measurements & aligned to the ASTM E3064-16 standard. This standard is relevant as IONA can dynamically track/measure; however, the practicalities of dynamic synchronisation between two systems are complex. The data acquisition of IONA is the same in static and dynamic operations, so it is justifiable to use the same dataset.
3. Absolute coordinate comparison: analogous to ASTM E2919-14.

Note: the IONA dataset has been reprocessed post-data collection to assess the latest improvements in the network adjustment.

4.1 Inter-point Distance VDI/VDE 2634-3 & ISO 10360-13 Analogue

This analysis is more aligned with the VDI/VDE 2634-3 & ISO 10360-13 standards and is analogous to ball beam measurements within the operational volume.

The analysis was performed by taking the inter-point distance of all 287 measurements within the grid, which equates to 40,000+ length measurements. The lengths range from 0.140 m to 1.715 m. Utilising the laser tracker measurements as the reference standard, we can calculate the measured difference and find the mean (bias) and standard deviation (variation) of the IONA measurements.

This analysis gives us a mean of 7µm and a standard deviation of 143µm with a maximum error of 700µm. Further removal of 0.1% of the outlying points gives a maximum error of 556µm. Outlier removal is a valuable sanity check to ensure that single outliers within our dataset are not adversely skewing the result.

4.2 Inter-point Distance ASTM E3064-16 Analogue

In order to align the analysis to the ASTM E3064-16 standard, only length measurements in the range from 250mm to 350mm were included from the total grid, resulting in 2000+ length measurements in total that fulfil this criterion.

This analysis gives us an essentially zero mean (0.4µm) and a standard deviation of 117µm with a maximum error of 700µm; further removal of 0.1% of the outlying points gives a maximum error of 532µm.

As a curiosity, if we apply a similar approach but using lengths over 1.25m, we yield a mean of 28μm and a standard deviation of 166μm with a maximum error of 658μm; further removal of 0.1% of the outlying points gives a maximum error of 540μm.

4.3 Absolute co-ordinate comparison ASTM E2919-14 Analogue

Although aligned with the ASTM E2919 standard, several departures were made for ease of data collection and comparison with the other standards. Rather than a 'random' selection of poses, the grid of poses was used. Also, because of the restrictions on the laser tracker reference measurements, the relative orientation of the robot end effector was unable to vary significantly throughout the data collection.

Absolute coordinate comparison compares the laser tracker measurements (our reference) to the IONA measurements from a common frame of reference (the fixture datum). The results in the table below express the mean (bias) and the standard deviation (variation) in each component, that is: x, y, z, from the reference measurement.

Note: the laser tracker/T-Mac orientation measurements (Rx, Ry, Rz) displayed a variation equal to that of the IONA system. As such, orientation has been excluded from the analysis as the laser tracker measurement cannot be said to be of a higher measurement quality than IONA and is unsuitable as a reference standard.

Table 1: Absolute co-ordinate comparison metrics

Statistic	**X (μm)**	**Y (μm)**	**Z (μm)**
Mean	-80	-230	-50
Standard Deviation	110	110	70
Abs Max Deviation	530	560	250

Although a valuable analysis for testing system performance, this methodology does – to a much greater extent than the relative analysis – conflate setup errors in establishing: i) a common reference frame and ii) a coincidence measurement frame between the two systems in line with ASTM E2919-14.

5 Conclusion

The absolute coordinate comparison analysis yields a result with standard deviations consistent with the relative analysis – as expected – however, there is a bias in the system, which is primarily missed when examining only the relative measurements. Most notably, this is in the y-axis, possibly due to a limited spread of reference targets along this axis within the volume. It is difficult to determine whether the setup has contributed to this observed bias. Whilst ASTM E2919-14 is possibly the most applicable standard available, the performance evaluation criteria are not in line with customer expectations for such evaluations, i.e., random poses throughout the volume.

The inter-point distance analysis gave a negligible bias, and the average system performance - standard deviation - was determined as ±143µm (using all data points). At this level of performance, there are several non-measurement error contributions. The setup of the common reference system is a complex error source to quantify that could be a significant portion of the measured deviations. Also, the single viewpoint laser tracker measurement performance of the reference measurements is not as good compared to the system under test as the standards would prefer.

This highlights the importance of datum system selection and understanding how best to design measurement processes. Arbitrary cell datums positioned large distances from the measurand are likely to magnify minor errors and lead to performance degradation. Local datum structures and relative moves should be used for higher performance.

References

[1] Verein Deutscher Ingenieure Verband Der Elektrotechnik Elektronik Informationstechnik (2002). *Optical 3D measuring systems Imaging systems with point-by-point probing*. (VDI/VDE no 2634-1)

[2] Verein Deutscher Ingenieure Verband Der Elektrotechnik Elektronik Informationstechnik (2012). *Optical 3-D measuring systems Optical systems based on area scanning* (VDI/VDE no 2634-2)

[3] Verein Deutscher Ingenieure Verband Der Elektrotechnik Elektronik Informationstechnik (2008). *Optical 3D-measuring systems Multiple view systems based on area scanning* (VDI/VDE no 2634-3)

[4] International Organization for Standardization (2021) *Geometrical product specifications (GPS) — Acceptance and reverification tests for coordinate measuring systems (CMS) — Part 13: Optical 3D CMS.* (ISO Standard no 10360-13:2021)

[5] ASTM International (2016) *Standard Test Method for Evaluating the Performance of Optical Tracking Systems that Measure Six Degrees of Freedom (6DOF) Pose* (ASTM no E3064-16)

[6] ASTM International (2014) *Standard Test Method for Evaluating the Performance of Optical Tracking Systems that Measure Six Degrees of Freedom (6DOF) Pose* (ASTM no ASTM E2919-14)

[7] Thompson, A., Southon, N., Fern, F., Stupfler, G., & Leach, R. (2021). *Efficient empirical determination of maximum permissible error in coordinate metrology*. Measurement Science and Technology, 32(10), 105013.

[8] Muelaner, J. E., Wang, Z., Jamshidi, J., Maropoulos, P. G., Mileham, A. M., Hughes, E. B. and Forbes, A. B. (2009), *'Study of the uncertainty of angle measurement for a rotary-laser automatic*

theodolite (R-LAT)', Proceedings of the Institution of Mechanical Engineers, Part B: Journal of Engineering Manufacture (223(B3)), 217– 229.

[9] Muelaner, J.E.; Wang, Z.; Martin, O.; Jamshidi, J. and Maropoulos, P.G. (2010). *Estimation of Uncertainty in Three-Dimensional Coordinate Measurement by Comparison with Calibrated Points*. Measurement Science and Technology, 21(2): pp.025106

[10] Martin, O.C., Robson, S., Kayani, A., Muelaner, J.E., Dhokia, V. and Maropoulos, P.G. (2016), *Comparative Performance between Two Photogrammetric Systems and a Reference Laser Tracker Network for Large-Volume Industrial Measurement.* Photogram Rec, 31: 348-360. https://doi.org/10.1111/phor.12154

[11] Hexagon Manufacturing Intelligence (2016). *Leica Absolute Tracker AT960: ASME B89.4.19-2006 Specifications*.

[12] International Organization for Standardization (1998). *Manipulating industrial robots — Performance criteria and related test methods* (ISO Standard no 9283:1998)