Evaluation of ICP variants for point cloud/BIM alignment enabling Scan-vs-BIM comparison: Application to maritime construction tolerance verification

Florie Becker [1], Hélène Macher [1], Tania Landes [1], Yvon Kila [2]

[1] Université de Strasbourg, CNRS, INSA Strasbourg, ICube Laboratory UMR 7357,
Photogrammetry and Geomatics Group, 67000, Strasbourg, France - (florie.becker, helene.macher, tania.landes)@insa-strasbourg.fr
[2] Ferrcad, 450 rue Baden Powell, 34 000 Montpellier, France - ferrcad@ferrcad.com

Keywords: Scan-vs-BIM, ICP, tolerance verification, maritime infrastructure, point cloud, BIM

Abstract

Reliable geometric verification is essential in the construction industry, particularly for large-scale maritime infrastructures where deviations can critically affect functionality and safety. The emerging *Scan-vs-BIM* approach enables automated quality assessment by comparing as-built point clouds with as-designed BIM models. It allows evaluation of the entire structure, rather than just specific points, but relies heavily on accurate spatial registration. This paper presents an evaluation of several Iterative Closest Point (ICP) variants for fine registration within a *Scan-vs-BIM* framework dedicated to construction tolerance verification. Three ICP variants are compared in terms of convergence behavior, robustness to noise, and stability using synthetic point clouds derived from maritime structures. The methods are then tested on real datasets, each acquired under different conditions, leading to varying data quality. Based on the results, a hybrid method is proposed to improve registration reliability. The results show that the proposed approach improves the inlier rate by 8–9 % while reducing the mean deviation by approximately 1 cm on the noisiest datasets, compared to the classical *point-to-plane* ICP.

1. Introduction

Errors in the construction domain have a significant impact on project costs, emphasizing the crucial need for reliable quality control. Traditional verification methods mainly focus on checking characteristic points of a structure, making them time-consuming and error-prone. As a result, they often overlook local deformations and cannot fully represent the actual structural condition. Recent technological advances, such as terrestrial, mobile laser scanning (TLS, MLS) and unmanned aerial systems (UAS), now enable the efficient acquisition of high-density geometric data with improved accuracy. Concurrently, Building Information Modeling (BIM) has become a key component of modern construction workflows, integrating geometric and semantic information within a shared digital model throughout the infrastructure lifecycle. The use of BIM facilitates collaboration among stakeholders by providing a shared, information-rich model throughout the infrastructure lifecycle. Compared to point clouds, BIM models offer structured data that can be exploited for control, verification, and monitoring during all project phases, from construction to operation and demolition. The comparison of laser scanning point clouds and BIM, known as *Scan-vs-BIM*, offers new opportunities for automated quality assessment and tolerance verification. Such comparisons enable the computation of point-to-model deviations, which can be exploited directly or to reconstruct building elements and to evaluate their dimensions, as well as to detect their presence or absence (Bosché et al., 2015) (Stilla and Xu, 2023). However, ensuring accurate spatial registration between as-built point clouds and as-designed BIM models remains a major challenge. The alignment accuracy directly affects the reliability of the resulting deviation analyses. Registration is generally the dominant contributor to the total measurement error, as any misalignment propagates directly to all subsequent deviation analyses. This step can be divided into coarse and fine registration. The former provides a rough alignment of the point cloud, independently of its initial position, while the latter refines this alignment by precisely adjusting the point cloud to a reference. Fine registration typically requires a good initial approximation to perform effectively. Consequently, achieving accurate registration is a critical step in the measurement chain. In practice, this task becomes increasingly challenging when dealing with large-scale point clouds containing millions of points, as registration can become computationally expensive. Additional difficulties arise when the object geometry is weakly constrained or when the point cloud is affected by significant measurement noise and outliers, which can degrade convergence and bias the registration results. Among existing fine registration methods, the Iterative Closest Point (ICP) algorithm and its variants are widely used, yet their performance depends strongly on the point cloud quality (noise, density, occlusions, initial pose) and the geometry of the scene. This paper presents an evaluation of ICP variants for construction tolerance verification of maritime infrastructures. In this study, both the point cloud and the BIM model are georeferenced within the same coordinate reference system (coarse registration), and only fine registration is subsequently studied. Section 2 introduces the need for fine registration in the context of tolerance verification. Section 3 presents the ICP algorithm and the selected variants considered in this study. Section 4 describes the experimental setup and evaluates the performance and behavior of the different variants on synthetic datasets. Section 5 proposes a hybrid approach combining the strengths of the most effective methods. Finally, Section 6 evaluates the performance of both the variants and the proposed method on real datasets.

2. Registration assessment for internal error determination

Tolerance verification in construction is commonly defined within national or international regulatory frameworks. In this study, the French decree of September 16th, 2003 (Ministère de l’Équipement, 2003) is used as a reference, as it provides explicit tolerance thresholds applicable to construction elements. While it does not explicitly address the *Scan-vs-BIM* process, similar principles are found in international standards such as ISO 19157, which define general frameworks for assessing the quality of geospatial data (Ariza-López et al., 2022). To ensure compliance with prescribed tolerances, deviations are computed between the as-built point cloud and the as-designed BIM model. Tolerances define acceptance thresholds, whereas measurement standard deviation characterizes the dispersion of repeated measurements and contributes to uncertainty estimation. Compliance is assessed by comparing the deviations to the tolerance thresholds, considering different confidence levels. The deviations between the point cloud and the BIM model can be decomposed into three main sources of error: internal error, legal reference network error, and georeferencing error. The latter two are related to external referencing processes and do not depend on the intrinsic quality of the point cloud. Therefore, this study focuses exclusively on the internal error, which reflects the intrinsic quality of the point cloud and includes effects related to sensor precision, data acquisition, and internal consistency after registration. It is evaluated by analyzing the deviations between the point cloud and the BIM model within an independent local coordinate system (Akca et al., 2010), after applying an optimal rigid transformation. This transformation corresponds to the registration step, where the ICP algorithm minimizes residual distances between corresponding points. As the internal error is assessed after registration, accurate registration is essential to avoid biasing the analysis. The selected ICP variants represent commonly used approaches for fine registration in industrial inspection workflows, assuming an approximate initial alignment. Global registration methods addressing coarse alignment, such as TEASER++ or KissMatcher, are therefore outside the scope of this study, as the datasets are already georeferenced.

3. ICP algorithm and variants selected for the study

The original variants of the ICP algorithm (Alg. 1) were introduced by (Besl and McKay, 1992) and (Chen and Medioni, 1992). Although both follow the same general principle, they differ in the error metric being minimized. The formulations are commonly referred to as *point-to-point* and *point-to-plane*, respectively. Since then, numerous variants have been proposed to improve robustness, accuracy and convergence behavior. In this study, three ICP variants were selected based on related work. The *point-to-plane* formulation is known to provide faster and more reliable convergence than the *point-to-point* version (Rusinkiewicz and Levoy, 2001). Moreover, the reference is a BIM model that can be represented as a mesh, only the *point-to-plane* family has been considered here. This formulation serves as the reference baseline for the comparison. The variants have been selected either for their known robustness to noise or for their ability to operate directly on meshes. The three selected variants are as follows :

1. **Point-to-plane ICP with a Cauchy robust kernel** (Babin et al., 2019). The squared residuals are replaced by a robust cost function $\rho(r_i)$, defined by Eq. 1 and 2.

$$E_{\text{Cauchy}}(R,t) = \sum_i \rho(r_i) \tag{1}$$

$$\rho(r_i) = \frac{c^2}{2}\,\log\Big(1 + (\frac{r_i}{c})^2\Big) \tag{2}$$

Algorithm 1 Iterative Closest Point (ICP)

Require: Source point cloud $P = \{p_i\}_{i=1}^N$, target point cloud $Q = \{q_j\}_{j=1}^M$, convergence threshold ϵ
Ensure: Estimated rigid transform (R,t) aligning P to Q

Initialize transformation $(R,t) \leftarrow (I,0)$
Set previous error $E_{\text{prev}} \leftarrow \infty$
repeat
  Correspondence: For each point $p_i \in P$, find its closest neighbor $q_i \in Q$ (using Euclidean distance or projection along surface normal)
  Transformation estimation:
    Compute the optimal rotation R and translation t minimizing the alignment error:
  - **Point-to-point:** minimize $\sum_i \|Rp_i + t - q_i\|^2$
  - **Point-to-plane:** minimize $\sum_i \big((Rp_i + t - q_i)\cdot n_i\big)^2$, where n_i is the normal at q_i

    The solution is obtained via least squares (e.g., using SVD or linearization).
  Apply update: Transform $P \leftarrow \{Rp_i + t\}$
  Check convergence:
    Compute the mean alignment error E and stop when $|E_{\text{prev}} - E| < \epsilon$.
    Set $E_{\text{prev}} \leftarrow E$.
until convergence
return (R,t)

The associated IRLS (Iteratively Reweighted Least Squares) weight is given by Eq. 3.

$$w_i = \frac{1}{1 + \left(\frac{r_i}{c}\right)^2} \tag{3}$$

This reduces the influence of points with large residuals.

2. **Point-to-plane ICP with an adaptive robust kernel (ARK-ICP)** (Chebrolu et al., 2020). Residuals are minimized using Barron’s generalized robust loss (Barron, 2017), whose shape parameter α_k is adapted at each iteration, defined by Eq 4.

$$E_{\text{ARK}}(R,t,\alpha_k) = \sum_i \rho(r_i(R,t),\, \alpha_k,\, c) \tag{4}$$

Barron’s robust kernel is given by Eq. 5.

$$\rho(r,\alpha,c) = \frac{|\alpha-2|}{\alpha}\left[\left(\frac{(r/c)^2}{|\alpha-2|} + 1\right)^{\alpha/2} - 1\right] \tag{5}$$

The shape parameter α_k is updated at each iteration based on the current residual distribution, typically by maximizing the likelihood of the residuals under Barron’s model. This adaptive scheme allows the loss to smoothly transition between different robustness regimes (e.g., from quadratic to heavy-tailed), improving resilience to outliers during early iterations while retaining accuracy near convergence. Barron’s robust kernel is similar to an earlier formulation of parametric robust loss functions, called Smooth Exponential Familly (Tarel et al., 2002), which already aimed at modeling several classical estimators such as Cauchy and Geman–McClure.

3. **Mesh-based ICP (MICP)** (Mock et al., 2024) (Mejia-Parra et al., 2019). Correspondences are obtained by closest-point projection onto the mesh via a raycasting acceleration structure, as shown in Eq. 6.

$$E_{\text{MICP}}(R,t) = \sum_i \left(n_i^\top \left(Rp_i + t - \Pi_{\text{CPC}}(Rp_i + t) \right) \right)^2 \tag{6}$$

where Π_{CPC} denotes the closest-point projection onto the mesh surface.

4. Evaluation of the variants

To identify the most suitable variant for the *Scan-vs-BIM* application, experiments were conducted to evaluate each method’s convergence basin, robustness to noise and stability. In this context, robustness refers to an algorithm’s ability to converge in the presence of noise, outliers, or incomplete data, whereas stability describes the consistency of the registration results when applying the same method to identical datasets under identical parameter settings. For the *point-to-plane* variants, the BIM model (used as the reference) must be sampled into a point cloud. It is important to note that the sampling strategy directly influences the results. If the sampled model contains too few points, correspondences between the two datasets may become unreliable. Therefore, it is recommended to sample the model with a density similar to or higher than that of the points cloud. Depending on the sampling method (uniform, Poisson, etc.), computation time may increase significantly. The influence of the sampling density was evaluated in preliminary experiments and kept consistent across all tests, as it is known to affect point-to-point distance computations (Lague et al., 2013). For the MICP variant, the results may be affected by the quality of the mesh, particularly if it contains holes, distorted faces or other defects. In the context of *Scan-vs-BIM*, however, the mesh used as the reference originates from a BIM model and can thus be assumed to be of high quality.

4.1 Synthetic datasets

The dataset used for the first tests is synthetic. A retaining wall without a coping beam made of concrete (Figure 1). This object represents a real-world maritime infrastructure element. Its well-constrained geometry and moderate size make it suitable for evaluating registration behavior under realistic conditions, while keeping the computational cost of the experiments reasonable. Using the BIM model of the infrastructure, a TLS acquisition was simulated from four stations to generate a point cloud (Figure 2). The simulation parameters were defined as follows, an angular resolution of approximately 0.25° horizontally and 0.30° vertically and a Gaussian noise with a standard deviation of 5 mm. The BIM model used as reference is converted into a point cloud using Poisson disk sampling (Yuksel, 2015).

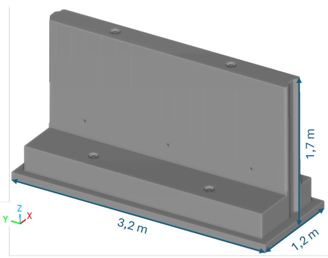

Figure 1. Infrastructure used for the experiments

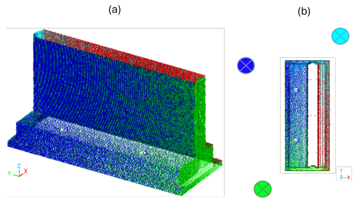

Figure 2. Infrastructure transformed into a point cloud (a) Point cloud created, colored by stations (b) Positions of the stations, one color per station

4.2 Robustness to the initial pose

Since the initial position of the point cloud is one of the main limitations of the ICP algorithm, this aspect was evaluated for each method. The convergence basin of an ICP method corresponds to the range of initial errors (rotation and/or translation) from which the algorithm is still able to converge toward the correct solution. A grid of initial transformations was defined, ranging from 0 to 30 cm for translations and from 0° to 30° for rotations, with a step of 2 units for each. To determine whether a method should be considered convergent, several convergence criteria were defined. Following the recommendations (Li et al., 2022), the first criterion requires the RMSE (Root Mean Squared Error) to remain below three times the standard deviation of the applied noise. Additional criteria were introduced to further ensure convergence: (1) RMSE must be below 1.5 cm; (2) at least 80% of the points must be classified as inliers (residual < 2 cm); (3) at least 95% of the points must have residuals below 3 cm. Firstly, rotation and translation were analyzed separately. For each test, an initial transformation was applied to the source point cloud according to the magnitude of the rotation or translation. For each step, 20 random directions were generated to define the transformation direction, resulting in 20 independent registrations per level. To ensure a fair comparison between the variants, no rejection threshold was applied, and an appropriate scale was adopted for the robust variants. For each method and each grid step, a success rate was computed as the proportion of tests satisfying all convergence conditions. These experiments (Figures 3 and 4) allow visualization of the stability domain of each ICP variant. A larger convergence basin reflects a better ability to recover from poor initial alignment, which is particularly important in real-world scenarios where precise initialization cannot be guaranteed, even when the datasets are georeferenced. Since real-world transformations typically involve simultaneous changes in translation and rotation, convergence was also evaluated using combined perturbations (Figure 5). For each configuration, only one trial registration has been performed. Combined with the separate tests, this analysis reveals which components of the transformation are more difficult for each ICP variant to handle. The ARK-ICP variant appears to be more sensitive to rotational errors, which is consistent with the findings of Chebrolu et al. (2020). The results highlight the behavior of the different variants clearly. The limits of the convergence basin depend strongly on the robust kernel used. In the case of ARK-ICP, the kernel becomes increasingly aggressive as the misalignment grows, classifying distant points as outliers and substantially reducing the convergence basin. With the Cauchy kernel, this effect is less pronounced, leading to slightly better convergence. The non-robust variants exhibit comparatively large convergence basins, which can be attributed to the absence of any distance-based rejection threshold typically applied in practice. It is important to note, however, that these results are facilitated by favorable geometry (strong constraints) and sufficient data quality. For more elongated or weakly constrained geometries, similar performance may not be achieved. Minor differences can be observed between the separate and combined experiments. This is due to the evaluation protocol.

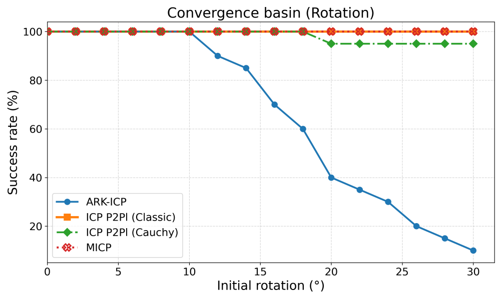

Figure 3. Convergence basin for rotational initialization errors, ICP P2Pl (Classic) and MICP curves are superimposed.

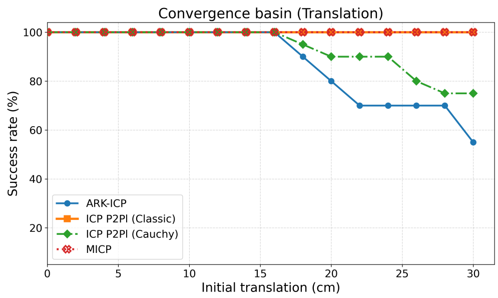

Figure 4. Convergence basin for translational initialization errors, ICP P2Pl (Classic) and MICP curves are superimposed.

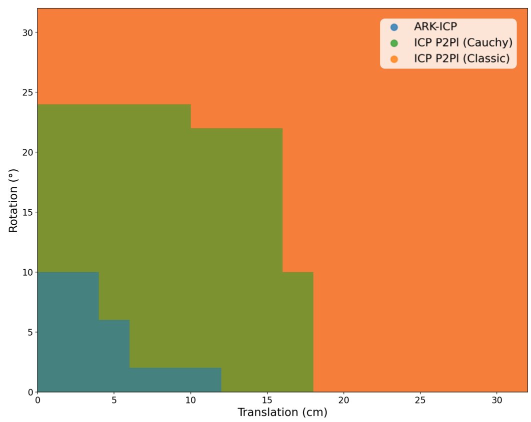

Figure 5. Convergence basin of each ICP variant, MICP covers the exact same area as ICP P2Pl (Classic).

4.3 Robustness to noise

Noise is unavoidable in real-world data. However, the concept of noise is not always clearly defined in the literature. In practice, three different phenomena can be considered as noise, although they have distinct origins and effects. Artefacts, corresponding to groups of points located far from the main point cloud, often result from reflections or sensor errors. Masks and occlusions can also be regarded as noise since they interfere with the completeness and quality of the acquired data. Finally, measurement noise refers to small random deviations introduced by the instrument itself, even when properly calibrated. Each acquisition device inherently generates a certain level of noise, and the characteristics of the scanned object can exacerbate it. Reflective or transparent materials, such as mirrors or water surfaces, tend to produce stronger disturbances. These observations emphasize the importance of data preprocessing before registration and the necessity to control every step of the processing pipeline to ensure sufficient data quality. Even when ICP is applied to a filtered or denoised point cloud, residual noise can remain and lead to poor alignment results. Therefore, it is essential that the fine registration method is robust to noise. In this section, only measurement noise was considered, occlusions have been considered in section 4.4. In the context of tolerance verification, it is assumed that the point cloud was acquired in accordance with state-of-the-art recommendations, ensuring sufficient quality and reducing occlusions. It is also assumed that the point cloud to be registered was preprocessed to improve its quality, with irrelevant areas and outliers removed. To quantitatively assess the influence of noise on registration performance, a dedicated experiment was conducted. The objective was to evaluate the robustness of each ICP variant when the source point cloud was affected by increasing levels of Gaussian noise. Ten noise levels were defined, ranging from 1 cm to 10 cm (standard deviation), distributed logarithmically to cover both low- and high-noise regimes. For each noise level, a small fixed transformation (a rotation of 1° and a translation of 3 cm) was applied to the point cloud. It was chosen to lie within the convergence basin of all considered variants, while remaining otherwise arbitrary, thereby enabling a fair comparison and isolating the effect of noise from that of the initial pose. Each ICP variant was executed on these noisy datasets. After registration, the estimated transformation (R,t) was compared to the known ground truth to compute two error metrics : the rotation error (°) (Figure 6) and the translation error (cm) (Figure 7). A rejecting threshold has been applied, fixed at 2 cm. It is commonly used to reduce the computational cost of the registration, ensuring that the algorithm will not search for correspondences too far away. Moreover, the rejecting threshold constitutes the sole means by which non-robust variants mitigate the influence of noise. The scale parameter of the Cauchy kernel and ARK-ICP was set based on the standard deviation of the applied Gaussian noise, following common practice in robust estimation (Chebrolu et al., 2020, Babin et al., 2019).

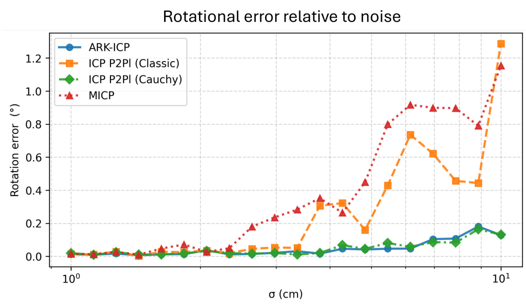

Figure 6. Rotation error as a function of the noise standard deviation (σ)

The results (Figures 6 and 7) highlight the instability of the MICP variant and the classic *point-to-plane* ICP when subjected to increasing noise levels. Both rotation and translation errors increase proportionally with the noise standard deviation, indicating a strong sensitivity to measurement perturbations and a lack of robustness. Conversely, ARK-ICP and Cauchy remain the most stable and consistent variants across all tested noise

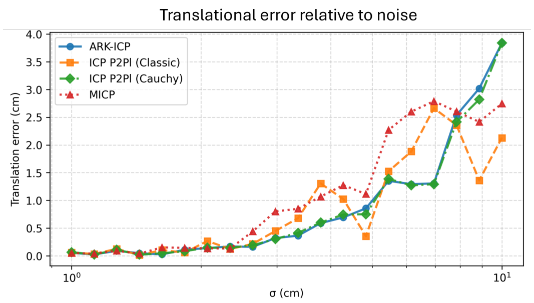

Figure 7. Translation error as a function of the noise standard deviation (σ)

levels, yielding nearly identical results. This behavior is expected, as Gaussian noise perturbs point positions and degrades correspondence estimation. The Cauchy kernel mitigates the influence of these mismatches by down-weighting large residuals, leading to improved robustness. Consequently, we observed that ARK-ICP selects a Cauchy kernel in these conditions, leading to similar performance.

4.4 Robustness to occlusions

Even when the point cloud is assumed to have been acquired in accordance with state-of-the-art recommendations, occlusions cannot always be avoided. A simulation of occlusions was performed at the scale of the entire structure to assess their impact on registration accuracy. The experiment was conducted on the BIM model of a boat lift. A model containing large additional objects was created to emulate masking effects during acquisition (Figure 8). Various elements were introduced, such as crates, floors, scaffolding, containers, and material pallets. The added elements were incorporated progressively, resulting in a gradual decrease in the proportion of inliers in the simulated point cloud. No preprocessing was applied, meaning that points corresponding to the added objects were retained. The applied transformation consists of a small translation mainly along the vertical axis (approximately 8.8 cm) and very small rotations (below 0.03°). This transformation was derived from the actual registration between the BIM model and the acquired real point cloud. The results show that occlusions have a limited effect on the translation error (Figure 9). Regarding rotation, no significant impact was observed, as the applied rotational perturbation was too small to produce measurable differences. Most variants exhibit an error of approximately one millimeter for an 8.8 cm translation, which remains acceptable. Even as the inlier percentage decreases, the error stays within the same range. However, the classic *point-to-plane* ICP demonstrates noticeable instability. This limited influence of occlusions was expected, since fine registration using ICP can be regarded as an optimization problem. Occlusions generally have a minor impact on optimization problem, whereas noise plays a significantly more influential role (Lee et al., 2020).

4.5 Stability of the results

As the instability of the classic *point-to-plane* variant has been demonstrated, an experiment was carried out to assess the stability of each ICP variant. Twenty registrations were performed on the boat lift dataset, using identical parameters (a rejection threshold of 2 cm and a robust scale of 0.005) and the same initial transformation applied at section 4.4. For each registration, the translation error was computed. The stability

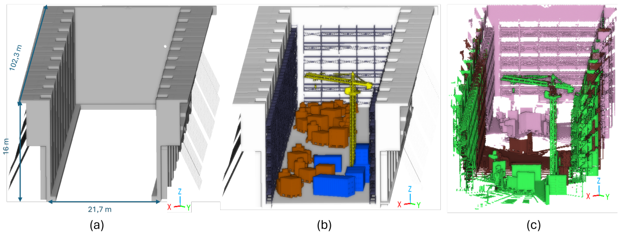

Figure 8. Masks simulation : (a) the original BIM model (b) the modified model (c) the simulated point cloud

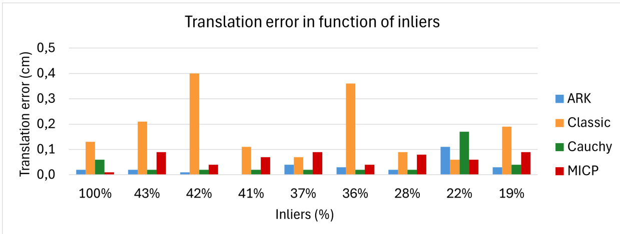

Figure 9. Translation error as a function of inliers

test reveals clear differences in the variability of the translation error across the ICP variants (Figure 10). The classic *point-to-plane* method exhibits the largest dispersion, confirming its sensitivity to small perturbations in the data and its tendency to converge inconsistently, even under identical conditions. The Cauchy-based variant shows moderate variability, indicating improved robustness but still a non-negligible dependence on local geometry or correspondence selection. ARK-ICP displays a markedly lower spread, reflecting stable convergence behavior and reliable optimization steps. Finally, MICP produces the most consistent results, with no observable variability across repetitions, demonstrating the determinism of the mesh-based correspondence strategy and its insensitivity to minor deviations in the input data. Overall, the comparison highlights that robust kernels and mesh-based formulations significantly improve the repeatability of the registration process compared with the classic *point-to-plane* ICP.

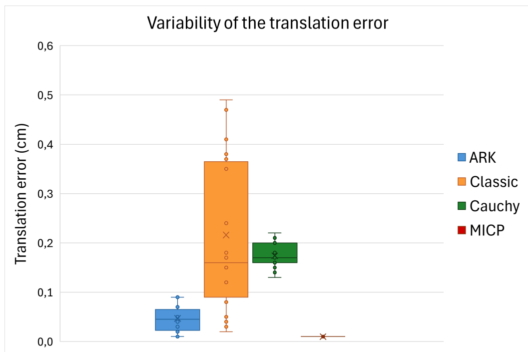

Figure 10. Variability of the translation error

4.6 Requirements for the *Scan-vs-BIM* pipeline

Regarding the data and the overall *Scan-vs-BIM* workflow, several criteria were defined to guide the selection of the most appropriate ICP variant. The specifications require that the *Scan-vs-BIM* pipeline must respect five conditions (1) robust to the initial pose; (2) robust to noise; (3) stable results; (4) no need to downsample the BIM model; (5) user-friendly configuration

This contribution has been peer-reviewed.
https://doi.org/10.5194/isprs-archives-XLIX-B2-2026-1229-2026

(minimal parameter tuning). The first three requirements are straightforward. However, the last two relate to practical considerations that strongly influence the reliability of the registration. It was decided not to downsample the BIM model in order to avoid incorrect correspondence selection and to eliminate the additional computational cost associated with this preprocessing step. Inappropriate parameter settings may lead to suboptimal results, even when a theoretically suitable method is used. Furthermore, the workflow must remain accessible to non-expert users, for whom identifying appropriate parameter values can be challenging. This may even be difficult for experienced users. The tests revealed both the limitations and the strengths of the different variants, summarized in the Table 1. As none of the variant satisfies all the requirements, a hybrid method is proposed.

Criterion	Classic	Cauchy	ARK-ICP	Mesh ICP
Downsampling req.	Yes	Yes	Yes	No
Robust to initial pose	Yes	Medium	No	Yes
Robust to noise	No	Yes	Yes	No
Stability of results	No	No	Medium	Yes
Requires settings	Yes	Yes	Yes	Yes

Table 1. Comparison of ICP variants according to *Scan-vs-BIM* requirements

5. Proposed hybrid method

The proposed approach combines the correspondence strategy of MICP, based on ray-casting onto the mesh surface, with the adaptive weighting mechanism of ARK-ICP (Alg. 2). While both methods individually show strong properties, their combination raises several theoretical and practical difficulties. First, the two formulations do not minimize the same type of residuals (Eq. 4 and Eq. 6), as seen at section 3. ARK-ICP operates on unsigned residual magnitudes, whereas MICP naturally provides signed *point-to-plane* distances, whose sign depends on the local surface normal. A difficulty arises from the geometry of the mesh. During the iterative registration process, the closest triangle associated with a point may change, resulting in abrupt variations of the surface normal. These normal flips induce discontinuous residuals, a residual can move from very small to very large between two iterations. Such discontinuities severely disturb the optimization of the kernel and therefore the robustness weighting. Another incompatibility concerns outlier rejection. MICP relies on a distance threshold, while ARK-ICP relies entirely on its weighting strategy and expects all residuals to be present. Combining the two strategies is possible but not efficient. When a rejection threshold is applied, the weighting strategy lacks sufficient residuals to adapt and identify an appropriate kernel. As a result, we observed that it tends to converge toward an L2 kernel, meaning that the weighting effect becomes negligible. In this configuration, the rejection threshold effectively suppresses the robustness provided by adaptive weighting. For this reason, no rejection threshold or annealing strategy was applied in this variant. The Hessian matrix associated with the *point-to-plane* formulation can also become ill-conditioned, especially in the presence of many coplanar correspondences, which is common in infrastructures. A normalization term is therefore added to stabilize the solver, which improves numerical behavior although it does not fully remove the intrinsic degeneracy. The normalization term $Z(\alpha)$ (also referred to as a partition function) ensures proper scaling of the loss and is precomputed numerically for a discrete set of α values. To address the incompatibility between signed MICP residuals and the unsigned nature required by ARK, the absolute value of the residuals is used for kernel evaluation and weighting. In addition, at each iteration, the loss is evaluated for all α values in a predefined grid and the α that minimizes the robust loss is selected. This adaptive mechanism significantly improves stability by allowing the method to adjust to the residual distribution at each iteration. These adjustments lead to a more stable and functional method, although at the cost of increased computation time compared to the classic ICP. Runtime optimization has not been addressed in this work, as several requirements had to be fulfilled first and computation time was considered secondary. This remains an avenue for future improvements. One of the design requirements of the proposed method was to keep the configuration as user-friendly as possible. Combining MICP with ARK-ICP introduces one parameter : the kernel scale c. The challenge is the selection of the kernel scale c which is known to be difficult. If c is too small, most weights collapse and the method becomes overly conservative, if c is too large, the kernel becomes almost quadratic, losing robustness. As explained by Chebrolu et al. (2020), it is theoretically possible to estimate the optimal pair (α, c) jointly. However, in practice, the two parameters are strongly correlated and tend to influence each other, making the joint optimization unstable. For this reason, Chebrolu et al. (2020) fixed the value of c before the optimization and the same choice is adopted in the current hybrid method. In theory, c should be chosen equal to the standard deviation of the noise in the point cloud. In practice, visually estimating this noise level is extremely difficult. To alleviate this issue, we propose a simple initialization strategy, before the first iteration, the distribution of point-to-surface residuals is analyzed and its empirical dispersion (using the MAD, mean absolute deviation) is used to compute an approximate value of c. This initialization does not aim at estimating the true noise standard deviation accurately, but rather at providing a reasonable and fully automatic scale parameter that improves the robustness and usability of the method. It should be noted that the estimation of c is only meaningful when the point cloud has been preprocessed so that non-relevant structures are removed. Points belonging to objects that are not present in the model would generate large residuals that do not correspond to sensor noise and would therefore strongly bias the scale estimation. Similarly, the estimation of c assumes that the initial pose is reasonably good. If the point cloud is located far from the mesh due to georeferencing errors or incorrect initialization (superior to 30 cm), the residual distribution will be dominated by a large global offset and the estimated c will be meaningless. In such cases, we recommend performing a coarse alignment first, for instance using RANSAC-based global registration, to obtain a sensible initial pose before applying the proposed method.

6. Tests on real data

To evaluate the performance of the tested variants under realistic conditions, four BIM/point-cloud pairs were selected, covering a wide range of geometries, scales, acquisition devices and noise levels (Table 2). For each dataset, five variants were compared: ARK-ICP, classic *point-to-plane*, *point-to-plane* with Cauchy kernel, MICP and the proposed hybrid method. Two indicators related to the registration process were computed : the overall RMSE (reported for completeness but not used for ranking due to fundamental differences in outlier handling between methods) and the percentage of inliers (points within the 2 cm tolerance). Four post-registration geometric deviation statistics were also computed : mean deviation, median deviation, standard deviation and the 95th percentile. Because robust-kernel variants retain all correspondences while

Algorithm 2 Hybrid ARK–MICP (Point-to-Mesh ICP)

Require: Source cloud $P = \{p_i\}$, target mesh $\mathcal{M}$, parameters
Ensure: Estimated rigid motion (R, t)

Initialize $R \leftarrow I_3, t \leftarrow 0$
Build RaycastingScene for $\mathcal{M}$ and precompute mesh normals
Precompute lookup table $\{\alpha_k, Z(\alpha_k)\}$
Raycast each p_i to obtain (q_i, n_i)
Compute signed distances $d_i = n_i^\top (p_i - q_i)$
Set robust scale $c = \mathrm{median}(|d_i - \mathrm{median}(d_i)|)$
for $k = 1$ to $K_{\max}$ **do**
Correspondences
  Transform $p_i^{(k)} = Rp_i + t$
  Raycast each $p_i^{(k)}$ to get (q_i, n_i) and compute $r_i = n_i^\top (p_i^{(k)} - q_i)$
Adaptive selection of α
  for each candidate α_k **do**

$$\mathcal{L}(\alpha_k) = \sum_{i=1}^{N} \rho(|r_i|, \alpha_k, c) + N \log\big(cZ(\alpha_k)\big)$$

  end for

$$\alpha^\star = \arg\min_{\alpha_k} \mathcal{L}(\alpha_k)$$

Computation of the weights
  Compute per-residual weights:

$$w_i = \frac{1}{|r_i|} \frac{\partial \rho(|r_i|, \alpha^\star, c)}{\partial r}$$

  Explicit expression:

$$w_i(\alpha^\star, c) = \frac{1}{c^2 \left(1 + \dfrac{r_i^2}{c^2 |\alpha^\star - 2|}\right)^{1-\alpha^\star/2}}$$

Linearized point-to-plane update
  Jacobian: $J_i = \left[((Rp_i + t) \times n_i)^\top \;\; n_i^\top\right]$
  Normal equations: $H = J^\top W J$, $g = J^\top W r$, with $W = \mathrm{diag}(w_i)$
  Regularize the system: $H \leftarrow H + \lambda I$
  Solve $H\,\delta x = -g$ for $\delta x = (\omega, v)$
SE(3) update
  $R \leftarrow \exp(\hat{\omega})R, \quad t \leftarrow \exp(\hat{\omega})t + v$, where $\hat{\omega}$ denotes the skew-symmetric matrix associated with ω
Stopping criterion
  Stop if RMSE or $\|\delta x\|$ is below thresholds or if $K_{\max}$ is reached
end for
return (R, t)

threshold-based methods explicitly reject outliers, the RMSE values are not directly comparable across families of methods. RMSE is therefore reported but not used as a decision criterion. Across all datasets (Tables 3, 4, 5 and 6), the hybrid method consistently achieves the highest inlier ratios while maintaining deviation statistics comparable to those of the robust ICP variants. These results indicate that the proposed approach is well suited to the *Scan-vs-BIM* context. It operates reliably across different geometries and acquisition conditions, remains practical to configure and delivers alignment quality on par with state-of-the-art robust methods. This confirms its relevance for tolerance verification workflows, where maximizing the proportion of well-aligned points is essential.

The real-data experiments reveal several consistent trends. Robust-kernel variants (Cauchy and ARK-ICP) generally produce stable deviation statistics, but their inlier ratios remain below those of the hybrid method. Classic *point-to-plane* ICP achieves low RMSE values due to its distance-based rejection, but its inlier ratio and deviation distribution indicate lower reliability in the presence of noise. Classic *point-to-plane* and MICP perform well on the boat lift dataset, where the point cloud is high quality, but their performance decreases on noisier datasets, highlighting the importance of robustness. The hybrid method systematically improves the inlier percentage, often by a significant margin, while preserving deviation metrics on par with the best robust variant. The hybrid method was also evaluated on synthetic datasets to analyze its behavior with respect to noise, convergence basin, occlusions, and result stability. The results are consistent with those obtained on real datasets, showing that the hybrid method combines the noise robustness of ARK-ICP with the stability of MICP. This confirms that combining mesh-based correspondences with adaptive weighting provides a strong balance between robustness to noise and geometric accuracy. Overall, the hybrid method meets the requirements of the *Scan-vs-BIM* pipeline more consistently than the other variants.

Site	Dimensions (m)	Sensor	# points
Dam	195 × 10 × 9	MLS	3 356 730
Boat lift	102 × 33 × 16	TLS	7 151 054
Technical room	5.9 × 3.5 × 2.3	GoPro	436 323
Technical room	5.9 × 3.5 × 2.3	iPhone	397 518

Table 2. Characteristics of the real datasets used for validation

Metric	ARK	Classic	Cauchy	MICP	Hybrid
RMSE (cm)	*8.3*	*0.9*	*8.3*	*1.1*	*8.3*
Inlier (%)	**32.6**	23.9	32.5	25.6	**32.6**
Mean (cm)	5.5	6.5	**5.4**	6.2	**5.4**
Median (cm)	**3.4**	4.7	**3.4**	4.4	**3.4**
Std. dev. (cm)	**6.3**	6.7	**6.3**	6.7	**6.3**
p95 (cm)	**17.9**	20.1	18.1	20.0	18.0

Table 3. Comparison of methods on the dataset of the dam

7. Conclusion

This study evaluated several ICP variants in the context of *Scan-vs-BIM* registration, considering robustness to initial misalignment, measurement noise, occlusions, and result stability. The analysis highlighted the respective strengths and limitations of classical ICP, robust-kernel formulations, and mesh-based approaches. Robust variants exhibited consistent behavior under noise, whereas classical *point-to-plane* ICP and MICP showed reduced reliability on the most challenging datasets. Conversely, non-robust variants were less sensitive to initialization when no rejection threshold was applied. Occlusions had a limited impact on the results, while noise remained the dominant factor affecting registration accuracy. The stability analysis further revealed that the classical *point-to-plane* formulation exhibited the highest variability, whereas mesh-based ICP benefited from a more stable correspondence strategy. These challenges are particularly critical in the context of large-scale infrastructures, where extended planar and coplanar surfaces (e.g., walls or decks) can lead to ambiguous correspondences and degraded registration performance. Overall, no existing variant fully satisfied the requirements of *Scan-vs-BIM* workflows. To address these limitations, a hybrid method was proposed, combining mesh-based correspondence selection with adaptive residual weighting. This approach was designed to meet key operational constraints: robustness to initialization and noise, minimal parameter tuning, and direct use of BIM meshes without downsampling. Experiments on real datasets showed that the

Metric	ARK	Classic	Cauchy	MICP	Hybrid
RMSE (cm)	*5.9*	*0.7*	*6.0*	*0.8*	*5.8*
Inlier (%)	81.2	81.8	81.9	**82.1**	81.8
Mean (cm)	**2.7**	**2.7**	**2.7**	**2.7**	**2.7**
Median (cm)	**0.8**	**0.8**	**0.8**	**0.8**	**0.8**
Std. dev. (cm)	**5.4**	5.5	5.5	5.5	5.5
p95 (cm)	**16.9**	17.0	17.0	17.0	17.0

Table 4. Comparison of methods on the dataset of the boat lift

Metric	ARK	Classic	Cauchy	MICP	Hybrid
RMSE (cm)	*7.9*	*0.8*	*8.0*	*1.0*	*7.9*
Inlier (%)	58.4	54.6	62.0	35.4	**62.4**
Mean (cm)	**4.3**	4.7	**4.3**	5.8	**4.3**
Median (cm)	1.5	1.6	**1.3**	3.6	**1.3**
Std. dev. (cm)	7.0	7.1	7.1	**6.3**	7.1
p95 (cm)	21.4	22.0	21.5	**20.7**	21.6

Table 5. Comparison of methods on the dataset of the technical room (Gopro)

proposed method improves the inlier ratio by approximately 8–9% compared to classical ICP, while maintaining deviation statistics comparable to robust ICP variants. These results demonstrate that the hybrid approach provides a reliable and practical solution for tolerance verification tasks, particularly in scenarios affected by noise, occlusions, or incomplete data. Beyond these results, this work highlights the importance of integrating additional sources of information into registration processes. In particular, the use of semantic information from BIM models could further improve robustness by guiding the registration toward structurally relevant elements (e.g., walls, beams) while reducing the influence of non-structural or transient objects. Finally, computation time was not addressed in this study. Future work will focus on optimizing the execution time of the hybrid ICP and exploring its integration into near-real-time quality control workflows. While the current method is designed for post-acquisition processing, such developments could extend its applicability toward more operational inspection pipelines.

Acknowledgment

The authors are grateful to Pierre Charbonnier (CEREMA/DTerEST/AS/ENDSUM) for his insightful suggestions and for his careful proofreading of the final version of this paper.

References

Akca, D., Freeman, M., Sargent, I., Gruen, A., 2010. Quality Assessment of 3D Building Data. *Photogrammetric Record*, 25(132), 339–355.

Ariza-López, F. J., Reinoso-Gordo, J. F., García-Balboa, J. L., Ariza-López, A., 2022. Quality specification and control of a point cloud from a TLS survey using ISO 19157 standard. *Automation in Construction*, 140, 104353.

Babin, P., Giguère, P., Pomerleau, F., 2019. Analysis of robust functions for registration algorithms. *2019 International Conference on Robotics and Automation (ICRA)*, 1451–1457.

Barron, J. T., 2017. A More General Robust Loss Function. *CoRR*, abs/1701.03077. http://arxiv.org/abs/1701.03077.

Besl, P. J., McKay, N. D., 1992. A Method for Registration of 3-D Shapes. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 14(2), 239–256.

Metric	ARK	Classic	Cauchy	MICP	Hybrid
RMSE (cm)	*5.3*	*0.9*	*5.4*	*1.0*	*5.4*
Inlier (%)	66.3	60.8	69.2	61.0	**70.1**
Mean (cm)	3.0	3.8	**2.9**	3.9	**2.9**
Median (cm)	1.1	1.3	**1.0**	1.4	**1.0**
Std. dev. (cm)	**4.5**	5.3	4.6	5.3	4.6
p95 (cm)	**13.2**	16.3	13.6	16.3	13.8

Table 6. Comparison of methods on the dataset of the technical room (Iphone)

Bosché, F., Ahmed, M., Turkan, Y., Haas, C. T., Haas, R., 2015. The Value of Integrating Scan-to-BIM and Scan-vs-BIM Techniques for Construction Monitoring Using Laser Scanning and BIM: The Case of Cylindrical MEP Components. *Automation in Construction*, 49, 201–213.

Chebrolu, N., Läbe, T., Vysotska, O., Behley, J., Stachniss, C., 2020. Adaptive robust kernels for non-linear least squares problems. *Proceedings of the IEEE International Conference on Robotics and Automation (ICRA)*, 4055–4061.

Chen, Y., Medioni, G., 1992. Object modelling by registration of multiple range images. *Image and Vision Computing*, 10(3), 145–155.

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, 10–26.

Lee, J. H., Park, J. J., Yoon, H., 2020. Automatic Bridge Design Parameter Extraction for Scan-to-BIM. *Applied Sciences*, 10(20), 7346. Number: 20 Publisher: Multidisciplinary Digital Publishing Institute.

Li, J., Hu, Q., Zhang, Y., Ai, M., 2022. Robust symmetric iterative closest point. *ISPRS Journal of Photogrammetry and Remote Sensing*, 185, 219–231.

Mejia-Parra, D., Lalinde-Pulido, J., Sánchez, J. R., Ruiz-Salguero, O., Posada, J., 2019. Perfect Spatial Hashing for Point-cloud-to-mesh Registration. D. Casas, A. Jarabo (eds), *Spanish Computer Graphics Conference (CEIG)*, The Eurographics Association.

Ministère de l’Équipement, 2003. Arrêté du 16 septembre 2003 relatif aux tolérances d’exécution des travaux de bâtiment et d’infrastructure. Journal Officiel de la République Française, JORF n°252 du 30 octobre 2003, texte n°18.

Mock, A., Pütz, S., Wiemann, T., Hertzberg, J., 2024. MICP-L: Mesh-based ICP for Robot Localization using Hardware-accelerated Ray Casting. *2024 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, 10664–10671.

Rusinkiewicz, S., Levoy, M., 2001. Efficient variants of the icp algorithm. *Proceedings Third International Conference on 3-D Digital Imaging and Modeling*, 145–152.

Stilla, U., Xu, Y., 2023. Change Detection of Urban Objects Using 3D Point Clouds: A Review. *ISPRS Journal of Photogrammetry and Remote Sensing*, 197, 228–255.

Tarel, J.-P., Ieng, s.-s., Charbonnier, P., 2002. Using robust estimation algorithms for tracking explicit curves. 2350, 492–507.

Yuksel, C., 2015. Sample Elimination for Generating Poisson Disk Sample Sets. *Computer Graphics Forum*.