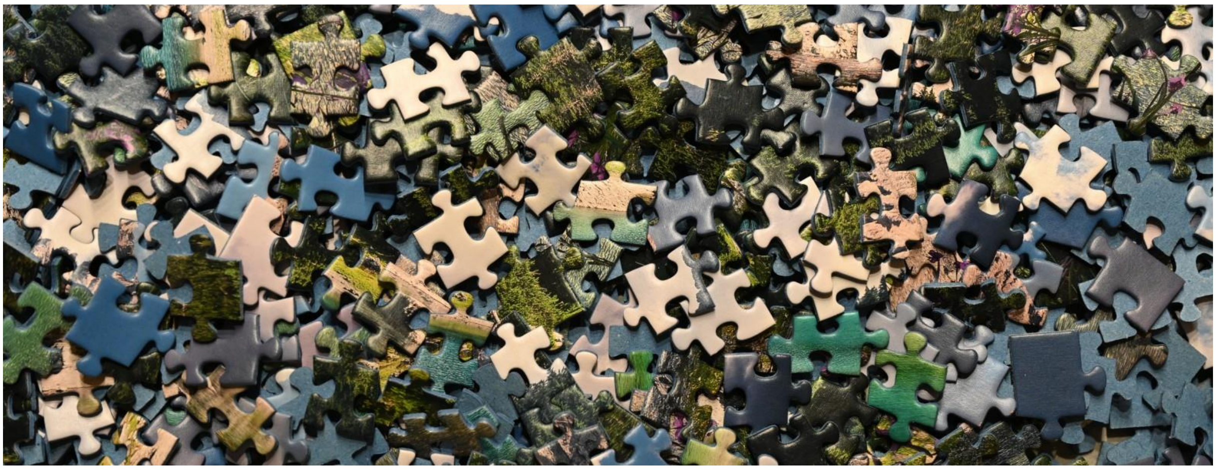
Taming errors... pt. 8: The difference between residuals, residuals and residuals

Daniel Wujanz – daniel.wujanz@technet-gmbh.com

25th March 2021



As mentioned in the very first episode of this series: overlooking registration errors can be quite painful in financial terms and additionally may leave deep dimples in your reputation. In the sixth episode we have also established that staring at the point cloud as a quality assurance procedure does not necessarily help to reveal all registration errors – especially smaller ones are hideous that typically become a problem once they accumulate. Thus, the focus of this issue is set on how errors can be found – regardless, which registration strategy was used and based on meaningful numbers. And yes, I am fully aware that "*meaningful numbers in laser scanning*" has not just a slight touch of an oxymoron to many of you.

A scientific field which has developed many expressive quality measures over the centuries is Geodesy / Surveying. This is why it is very surprising that the vast majority of the entire laser scanning industry is associating "quality" with just a single measure – namely residuals. The foremost problem is that residuals can be computed in different ways. Thus, you have to understand what type of residuals you are looking at, what they tell you and, most importantly, what they do not express. In one of the next episodes, we will also have a look at additional QA metrics and how they can help you to sleep a lot tighter.

So, why are residuals so popular? Well, residuals are, in geodetic terms, *datum-independent*. That means that these quality measures always remain the same, even if you translated your point clouds beyond mars or rotated them upside down. The first problem is that residuals can refer to a) pairwise registrations where residuals between two scans are minimised (Besl & McKay 1992) or b) residuals that stem from a block adjustment (Pulli 1999) where inevitable discrepancies among all redundant pairwise registrations are minimised. To make it short: always use the latter since residuals among pairwise registrations typically shine like gold yet quite often turn out to be something very different once they are assessed by the mighty sword of redundancy. The biggest problem is however that there are three strategies to compute residuals between point clouds after registration, namely,

- Residuals between discrete points (target centres)
- Residuals between the point clouds itself
- Residuals between redundant registration parameters (see part 7 and the next issue)

Artificial targets are the most trusted registration strategy in practice due to the simple fact that their quality measures are expressive and easy to interpret. Since the general workflow and metric interpretation generally mimics working with tacheometric observations these numbers feel quite familiar for many people, particularly for surveyors. Figure 1 illustrates a scenario with three stations that were registered with three spheres. The colours of the spheres highlight by which scanner they were captured. Since the triplet of spheres was captured by three stations, nine spheres can be found in the figure. Even though artificial targets, such as spheres, allow to compute discrete target centres, these are never error free due to measurement noise and the target fitting error. The latter error originates from the computation of a target's centre point which can either stem from e.g., correlation-based approaches, that are typically used for checkerboard targets, or an approximation of points that were recorded on top of a spherical target. As for the given example a block adjustment with six point-to-point-correspondences was computed, based on which nine residuals can be calculated. The residuals are the remaining tensions after the adjustment, which are also referred to as misclosures in surveying, between the individual target centres. Hence, for each one of the three spheres you will receive residuals in the colour code red and blue, red and green as well as blue and green.

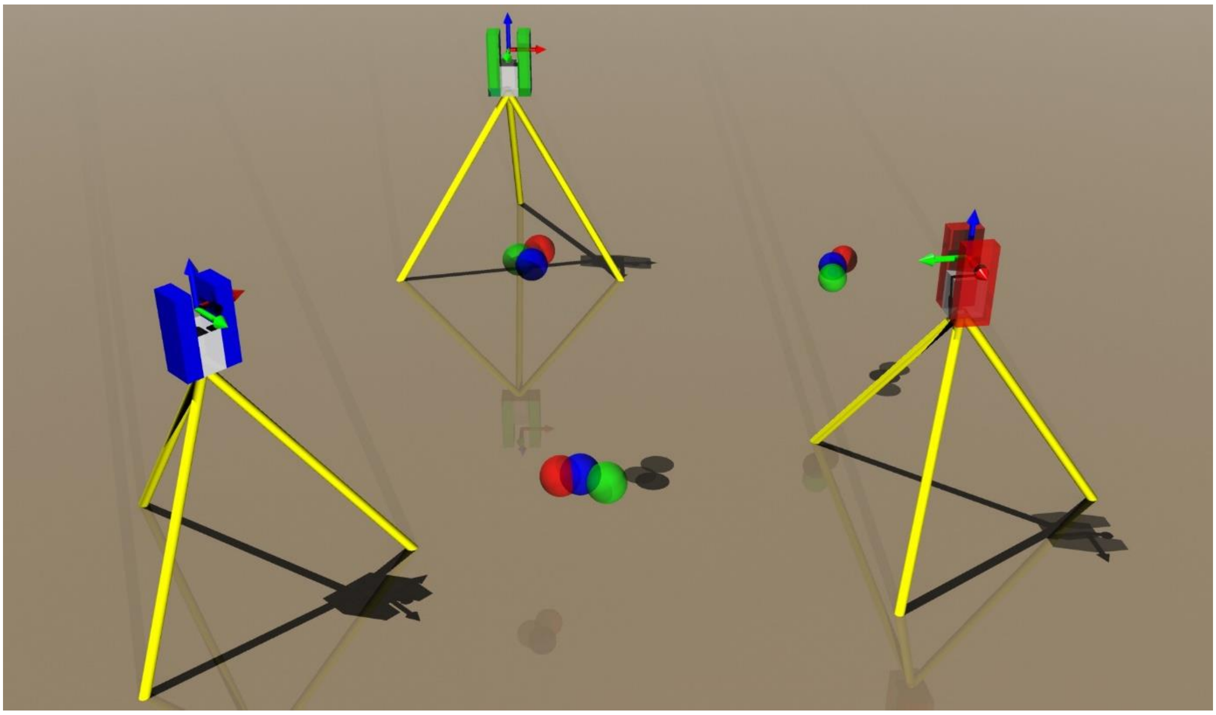


Figure 1: Three stations registered with three spheres

The reason why the vast majority of people in the laser scanning business appear to severely suffer from Tourette's is strongly correlated to the second interpretation of residuals: geometric differences between two point clouds. The first issue is that most users are not aware how correspondences are formed (see Taming error pt. 2) – which means that identical registration parameters lead to different quality measures. Let us consider two registered point clouds that are highlighted by orange and green spheres in Figure 2. The observed surface is highlighted by the dark grey line. Note that is NOT important how the point clouds were registered. Four green spheres were selected in order to compute the quality measures for this registration. The chosen correspondence strategy is a point-to-point-based approach as highlighted by the yellow links. Their lengths represent the size of individual residuals. As a final step a single numerical value is computed, e.g., the mean value of all residuals. For the given case, this value sums up to 7.8 mm on average (residuals from left to right: 7.8 mm, 7.0 mm, 7.4 mm, and 9.0 mm).

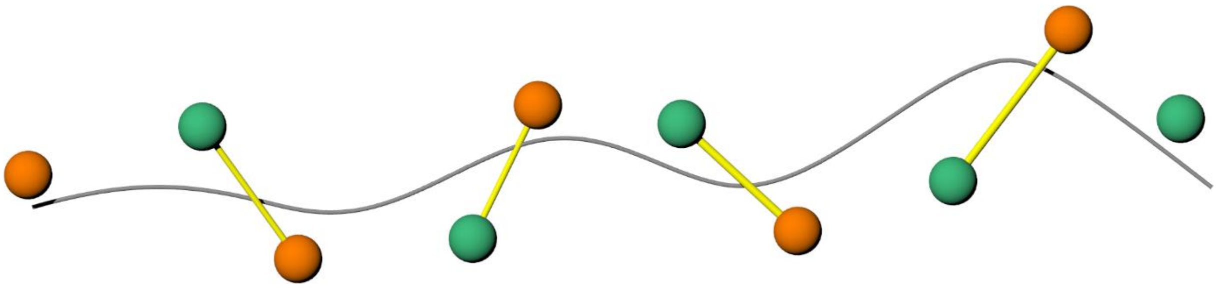


Figure 2: Slice of registered point clouds illustrating point-to-point metrics.

It is obvious that the impact of local sampling has a major impact onto the previously computed quality measures. Hence, the identical scenario shown in Figure 2 is now being processed by calculating point-to-triangle correspondences as depicted in Figure 3. As a first step, the orange dataset was triangulated as highlighted by the straight segments tinted in light grey. Subsequently, the green points are projected onto corresponding triangles. The distance between a green point and the projected base point represents the resulting residual. In this example the “quality” of the registration is 3.4 mm on average (residuals from left to right: 5 mm, 3.5 mm, 2.6 mm, and 2.5 mm). By solely looking at the numbers one could assume that the quality of this registration would be ~twice as accurate as the first one – even though the registration is identical.

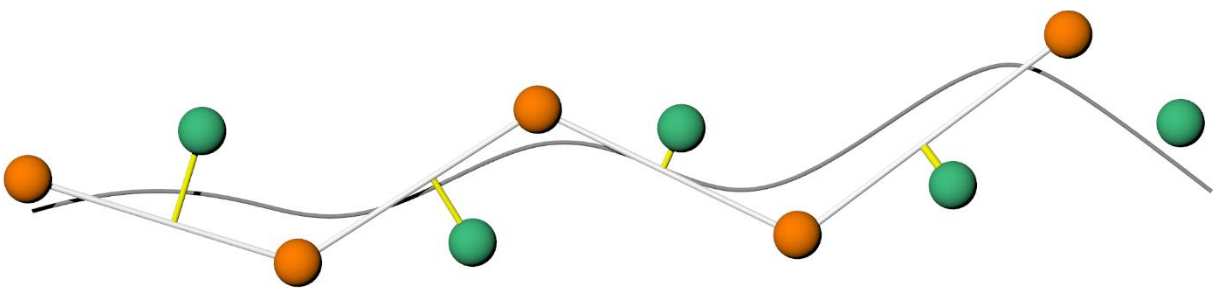


Figure 3: Slice of registered point clouds illustrating point-to-triangle metrics.

In the next example, which is illustrated in Figure 4, the green dataset was rotated and translated with respect to its original location in order to demonstrate the impact of a subtle parameter. Sometimes this value goes by the name of e.g., maximum search distance (for some cloud-2-cloud implementations), correspondence threshold or tolerance and is sometimes tuneable in the properties of the applied registration algorithm. This parameter specifies the largest distance respectively residual between two points from two different point clouds that can form a correspondence. In the following this parameter is considered to be 3.0 mm. The residuals in this example (from left to right) are 12.0 mm, 2.6 mm, 2.9 mm, and 12.7 mm. Two residuals are rejected since they are larger than the specified threshold, as highlighted by the two red links. Consequently, the “quality” of the registration is 2.75 mm which does not sound bad at all. If we now apply the same parameter to the example depicted in Figure 3, the quality measures “improve” from 3.4 mm to 2.55 mm which is of course irritating. In brief: if this distance is set to 1 mm, the resulting quality measure will be smaller than 1 mm.

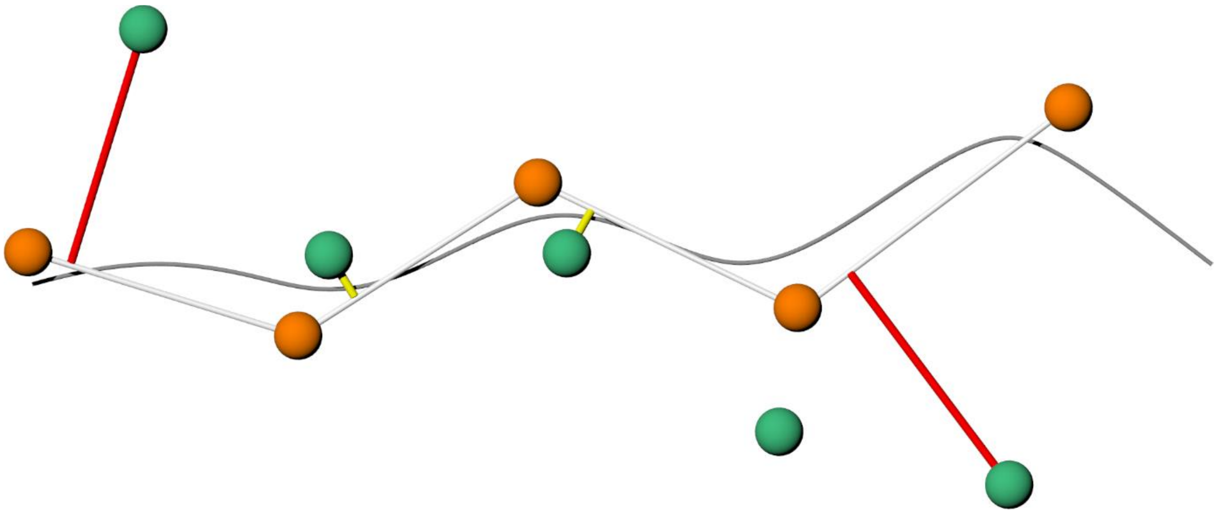


Figure 4: Slice of registered point clouds illustrating point-to-triangle metrics with rejection of residuals with a certain magnitude.

In the previous examples we computed mean residuals that serve as quality measures. Yet, sometimes also the median is reported, or the value considers only the best 83.7518% (or what not) residuals. The biggest issue is however that nobody knows which points were used! It could well be that the metrics are reported for correspondences in a very small overlapping region between the registered datasets (see Figure 6 of Taming error pt. 2 or Figure 4) with great looking numbers despite a rubbish outcome. Having no information which points were used also means, and I repeat myself, that identical registration parameters can lead to different quality measures. Even if we could visualise which points were used to compute the residuals (Wujanz 2012) it would only be feasible to use in the case of having to check just two scans opposed to thousands of scans where other quality make sense. Finally, it has to be assumed that the reason why people are ranting and raging against cloud-to-cloud registration even though residuals between the point clouds itself as a quality measure need to be blamed. Hence, it is rather disturbing that the vast majority of an entire industry and even science still uses this metric.

References

Besl, P. J., McKay, N. D. (1992). Method for registration of 3-D shapes. Robotics-DL tentative, International Society for Optics and Photonics, pp. 586-606.

Pulli, K. (1999): Multiview registration for large data sets. In Second International Conference on 3-D Digital Imaging and Modeling (Cat. No. PR00062) (pp. 160-168). IEEE.

Wujanz, D. (2012): Towards transparent quality measures in surface based registration processes: Effects of deformation onto commercial and scientific implementations. Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci., Vol. 39, Part B5, pp. 251-256.