Invalid measurement data are phenomena that plague the reliability of Coordinate Measurement Systems. The best-fit algorithms used for feature computation all amplify the influence of the points most distant from the mean. As a result, a single invalid point can significantly skew the computed feature parameters, even for large data sets. If the feature happens to be a datum, the error propagates through the measurement of the entire part. A measurement consisting of thousands of measured points can be ruined by a single bad piece of data. It is therefore desirable that invalid data be identified and removed before feature evaluation is performed.

Three separate cases are considered. The first case is the problem of contact occurring on undesired surfaces. This type of problem can occur when measuring complex parts such as a gear hob or when using a non-spherical probe. This case is the best suited for solution through the use of surface normal data since the change in surface normal between surfaces is large and easily detected. The remaining problem, the detection of invalid data due to dirt, burrs, porosity or external factors, is divided into 2 cases. The division is made based upon the spacing of the data collected. If the spacing is large, each point can only be considered independently. If the measurement data are dense we can use the relationship between points to refine the analysis.

**Complex Part Evaluation**

Complex parts can be vulnerable to invalid data due to inadvertent contact with nearby surfaces that are not part of the data set. Cutting tools such as broaches or gear hobs are a particular instance of this problem. The teeth are closely spaced and each tooth is composed of a number of surfaces that need to be evaluated separately. Due to the small clearances between surfaces, it is common for a measurement that was targeting one surface to actually collect data from an adjoining surface. A small error in the rotational alignment used to collect the measurements can lead to large numbers of data points being attributed to incorrect features. A similar problem occurs when non-spherical probes are used for measurements. Figure 1 shows three types of non-spherical probes commonly used for coordinate measurement. These probes have a limited set of surfaces that are valid for metrology purposes. It is desirable to be able to recognize when data are collected from contact with invalid parts of the probe.

In both of these cases there is a large variation in surface normal values between each surface. The cutting tool example has surfaces that vary in direction by greater than 10 degrees. If the surface normal at each point is measured it is a simple matter to identify the actual contact surface for each point. The process to properly associate the measured data with each feature would be a two step process. First, the points would be separated into groups related to each tooth surface based upon the surface normal measured. Then the sets would be associated with the appropriate tooth number based upon the location of points in space.
A similar approach could be applied for measurements using non-spherical probes of the type shown in Figure 1. Measurements collected using these probes can be divided into 3 categories based upon the value of the measured surface normal. The categories are valid 3D data, valid 2D data and invalid data. Consider the case of a cylinder sphere probe. The relevant categories as a function of the angle the measured surface normal makes with the axis of the probe cylinder would be:

<table>
<thead>
<tr>
<th>Angle to Axis</th>
<th>Case</th>
</tr>
</thead>
<tbody>
<tr>
<td>0° to 89.x°</td>
<td>Valid 3D data</td>
</tr>
<tr>
<td>90°</td>
<td>Valid 2D data</td>
</tr>
<tr>
<td>&gt;90°</td>
<td>Invalid data</td>
</tr>
</tbody>
</table>

The transition value between 3D and 2D data would be found by subtracting the surface normal measurement repeatability from 90°.

**Sparse Data Sets**

With a sparse data set, it is not possible to make a comparison between adjacent points. There are still a number of invalid data conditions that can be detected by simply examining the measured surface normal of each point. Interference from burrs or incorrect surface contact, vibration triggering of the probe and large pieces of foreign material such as machining chips all produce large surface normal deviations. These kinds of probing errors can be detected and excluded by selecting a heuristic threshold value for the deviation of the surface normal from nominal. The threshold test would be applied after a surface fit is performed to remove the potential of a complete data set being flagged due to a tilt error in manufacturing. If any invalid points are detected and removed the fit would be computed a second time without the invalid

![Figure 1: Profiles of commonly used non-spherical probes. Figure a) represents a cylinder-sphere probe, b) is a disk probe and c) is a point stylus](image-url)
The size of such a threshold would vary depending upon the surface condition of the part under examination.

The objective is to select a threshold value sufficiently small to exclude invalid data but large enough to ensure a small probability of rejecting valid results. If the process accepts invalid data, the result will be a false report suggesting an out of tolerance condition. The cost of such a report at worst some lost production time while the error is investigated and, in most cases, the only cost will be the cost of a repeated inspection of the part in question. However a valid point flagged as invalid could potentially result in a false report indicated that an unacceptable part meets the design standard. The potential cost of this type of error is the production of numerous defective parts that will need to be scrapped at some future point in the manufacturing process. In the extreme, the production of defective parts could lead to the loss of a customer or insolvency of the producing company. The asymmetrical nature of this problem makes it clear that any choice of threshold value should be biased to ensure there is a negligible probability of rejecting valid data.

To select a suitable threshold value, an understanding of the magnitude of angular errors present on the surface is required. The magnitude is a function of 2 factors: the amplitude and wavelength of the surface error. Increasing the size of the surface errors results in larger angular errors. Smaller wavelengths produce larger angular errors as well. Amplitude and wavelength are not strictly independent since energy considerations tend to dictate that smaller wavelengths have relatively smaller amplitudes. However, for the purpose of selecting a threshold value, amplitude and wavelength will be considered to be independent. The result will be a slightly more conservative test value, but this satisfies the intent of biasing the threshold against rejecting valid data.

If the surface error is assumed to be a single sinusoid of wavelength L and amplitude A, the maximum angular error on the surface is:

$$\tan^{-1}\left(\frac{A}{L}\right)$$

The amplitude A should be set to ½ the largest error that can be reasonably expected from the manufacturing process for the surface in question. The wavelength parameter should be the width of the smallest deviation that is desired to be detected by the measurement process. A reasonable choice for this value would be the sample spacing. As an example using 0.050 mm for the error amplitude and 5 mm as the wavelength parameter yields an angular threshold value of 3.6°.

### 8.4 Dense Data Sets

In the case of dense data, the relation between adjacent points can be used to form a frequency filter to detect invalid data. Since the concern is with the exclusion of outlying points in the data set, the computational load can be reduced by only considering points that are in the outer quartiles of the total range of deviations in the set. The method works by comparing the error in the surface normal of the surface with the expected value of the second derivative of the surface based upon the assumption that the surface is locally smooth and continuous. If the surface normal error exceeds the second derivative criteria, the data point is flagged as invalid.

Consider the 3 points shown in Figure 2, $d_j$ is the deviation from nominal for the jth point, $\theta_2$ is the tangent error angle found from measuring the surface normal at point 2, and $s$ is the distance between sample points. Presuming a surface shape that approximates a sinusoid, the second derivative of the surface would be of the form:

$$\frac{\partial^2 x^2}{\partial^2 y} = A^2\left(\frac{2\pi}{L}\right)^2 \sin\left(\frac{2\pi}{L} x\right)$$
To select suitable values for the amplitude of the surface $A$, and the wavelength, $L$ we consider the measurement parameters. Since we have made the assumption that our data set includes the full range of deviations present on the part it is safe to set the value of $A$ to be the range of deviations present in the set of three points under consideration. The wavelength parameter is set to be the distance across the group of points, $2s$. When $s$ is particularly small, $L$ may be set to be either of the minimum feature size based upon mechanical filtering of the probe sphere or a minimum wavelength cutoff specified by the user. Using this model, the maximum second derivative we would expect to encounter on a smooth continuous surface would be:

$$\frac{\partial^2 x^2}{\partial y^2} = \left(\text{Range}[d_1, d_2, d_3]\right)^2 \left(\frac{2\pi}{2s}\right)^2$$

If the value of $\varepsilon$ calculated above exceeds this test criteria, we can conclude that the point in question is not representative of a smooth, continuous, surface and should be excluded from the data set.

**Acknowledgements:**

This work has benefited greatly from the exceptional research environment provided by the Center for Precision Metrology at UNC Charlotte and the Cameron Center for Applied Research. This material is based upon work supported by the National Science Foundation, the UNC Charlotte Center for Precision Metrology Research Affiliates and the National Science and Engineering Research Council of Canada.

**References:**
