

Explanations for Using the NIST Reference Pairs: Data Sets and their Least-Squares Fits

The zip file, “NIST-L2-reference-pairs,” contains eight directories with a total of 240 data sets along with their corresponding least-squares fits. The directory names identify the geometry: Line, Plane, Circle2d, Circle3d, Sphere, Cylinder, and Cone. Every data set name ends with the extension “.ds” and its corresponding fit has the same name but ends with “.fit”

Generally, the data sets were generated according to the ASME B89.4.10-2000 standard (incorporating the 2002 correction on aspect ratios). The default test was used, which means the sizes of parts range from 1 to 500 (mm, but the units are not essential), the farthest position of the part in the coordinate system is 1000 mm, the maximum number of points in a dataset is 500 and the coordinate values are given with five digits to the right of the decimal place.

These reference pairs are provided for industrial use for convenient self-checking. As these data sets are given along with their reference fits, they should only be used for internal “practice” testing and not in lieu of testing by an independent body like NIST.

Even though the data points are given with this fixed precision, the mathematical problem assumes an arbitrary number of trailing zeros, so the reference fit files are given to greater precision than the data appear to have. We believe the reference fits to be correct to all the digits we report.

The data sets are plain text files, each having its first line indicating the number of points in the data set and each line following having three coordinates. An example of the first four lines of a file looks like:

```
72
85.0672      13.01812    19.23795
85.00934    12.92935    19.2457
84.98234    12.87145    19.24716
```

The data sets for the two-dimensional geometries are restricted to a plane parallel to a coordinate plane—not necessarily the x-y plane. The data are still given in x-, y-, and z-coordinates. (So in these cases, one of the three coordinate values will be the same for all points).

The data sets have no special point ordering or “seed values” that some algorithms require. They also include partial features where applicable and where the standard specifies them.

A fit file contains a list of numbers that define the least-squares fitting geometry. The number of points and their meaning is as follows:

Lines – 6 numbers

3 numbers represent a point on the line

3 numbers represent the direction cosines of the line

Planes – 6 numbers

3 numbers represent a point on the plane

3 numbers represent the direction cosines of the normal to the plane

Circles – 7 numbers

3 numbers represent the center of the circle

3 numbers represent the direction cosines of the normal of the plane containing the circle

1 number represents the diameter of the circle

Spheres – 4 numbers

3 numbers represent the center of the sphere

1 number represents the diameter of the sphere

Cylinders – 7 numbers

3 numbers represent a point on the cylinder axis

3 numbers represent the direction cosines of the cylinder axis

1 number represents the diameter of the cylinder

Cones – 8 numbers

3 numbers represent a point on the cone axis

3 numbers represent the direction cosines of the cone axis

1 number represents the orthogonal distance from the reported point on the axis to the surface of the cone.

1 number represents the full apex angle of the cone in degrees (less than 180)

A note on three-dimensional circles: There are two approaches to their least-squares fitting: 1) Fit the data to a least-squares plane, project the data onto that plane, then fit a circle to the projected points in a least-squares sense, or 2) perform a full-3d least-squares fit to the data, defining each residual distance as the shortest 3d distance from the data point to the curve. The fits here were done according to the first strategy, which is by far the most commonly used and in most measuring situations is the correct approach to use. We can also provide reference fits for the second method if needed.

A note about comparing other fit results to these reference fits: In the case of a cylinder, for example, a reference fit would include a point on the axis, the direction cosines of the axis, and the diameter. Two representations of the same cylinder might use different points on the axis, so two solutions may appear different when they are the same. The diameter comparison would not have this problem, nor would the comparison of direction cosines (up to the 180 degree rotation of the axis direction.)

While the direction of the cylinder axis can be reported with a 180-degree ambiguity, this is not the case with the direction of the cone axis. There should be something in the representation of the cone to not only indicate the axis direction, but the direction of

opening with respect to that direction. In these reference results the reported direction, when taken from the reported point on the axis, always points in the direction of increasing radius of the cone. (In the reference results the vertex is never used as the reported point on the axis.)

In the one case of a cone where the data set contains six points, it might be possible that the least-squares solution is not unique. Do not be alarmed if you are comparing results and find significant differences in this one case, as the difference is probably due to the problem of nonuniqueness.